The Recommended GHRSST-PP Data Processing Specification GDS (Version 1 revision 1.6)

GHRSST-PP Report Number 17.

Reference:	GHRSST/17
Issue:	GDSv1 Revision 1.6
Date of issue:	Wednesday, 20 April 2005
Document type:	Reference document
Author:	The GHRSST-PP Science Team and GDS-TAG working group

The Recommended GHRSST-PP Data Processing Specification GDS (Version 1 revision 1.6)

GHRSST-PP Report Number 17.

Compiled by Craig Donlon and the GHRSST-PP Science Team
craig.donlon@metoffice.gov.uk

Published by the International GHRSST-PP Project Office
Met Office
Fitzroy Road
Exeter, EX3 1PB
United Kingdom

Tel: +44 (0)1392 886622
Fax: +44 (0)1393 885681

The GHRSST-PP International Science Team (2004/05):

Craig Donlon	Chair, Hadley Centre, Met Office, United Kingdom
Bill Emery	University of Colorado, USA
Hiroshi Kawamura	Tohoku University/JAXA; Japan
Jim Cummings	Naval Research Laboratory, Monterrey (Ca), USA
Ian Robinson	Southampton Oceanography Centre, United Kingdom
Pierre LeBorgne	Meteo France/O&SI SAF, France
Doug May	Fleet Numerical Met. and Oceanog (NAVOCEANO), USA
Peter Minnett	University of Miami, USA
Ian Barton	CSIRO Marine, Australia
Nick Rayner	Hadley Centre, Met Office, United Kingdom
Chelle Gentemann	Remote Sensing Systems, USA
Chris Mutlow	Rutherford Appleton Laboratory, United Kingdom
Gary Wick	NOAA ETL, USA
Andrew Harris	NOAA/NESDIS ORA, USA
Ed Armstrong	JPL PO.DAAC, USA
Jorge Vasquez	JPL PO.DAAC, USA
Kenneth S. Casey	NOAA/NESDIS NODC, USA
Richard Reynolds	NOAA NCDC, USA
Bob Evans	University of Miami, USA
David Llewellyn-Jones	University of Leicester, UK.
Jean-Francois Piolle	IFREMER, Brest, France.
Andrew Bingham	JPL PO.DAAC, USA
Chris Merchant	University of Edinburgh, UK.

The following individuals have provided significant inputs to the preparation and review of the GDS:

Alice Stuart-Menteth	Southampton Oceanography Centre, United Kingdom

LEGAL NOTICE

Neither the Met Office nor any person acting on behalf of the Met Office is responsible for the use that might be made of the following information.

Printed in England.

Document change record

Title:
The Recommended GHRSST-PP Data Processing Specification (GDS)

GHRSST-PP reference document number:
GHRSST/17

Approved by:

Name	Entity	Signature.	Date
G Wick	For the GDS-TAG	Gary A. Wick	04/01/2004
C J Donlon	For the GHRSST-PP Science Team		23rd March 2004

Document change record:

Author	Modification	Issue	Rev.	Date
C Donlon	Basic v3.0 complete.	DRAFT	V0.1	10th September 2002
C Donlon	Medspiration 1.5d URD and GDIP 1.2. now merged into GDS following discussions with ESA at GHRSST-III. Included significant changes based on the GHRSST-PP workshop which resulted in a complete re write and specification of the GDS.	DRAFT	V0.2	22nd December 2002
C Donlon	Added J. Vasquez and E. Armstrong comments on data management, MMR and data exchange. Comments from K Casey on data formats and QC	DRAFT	0.3	31st January 2003
C Donlon	Added general comments from I. Robinson	DRAFT	0.3.1	8th February 2003
C Donlon	General inputs from data providers added following January blanket mail request for comments	DRAFT	0.3.5	10th February 2003
C Donlon	Added section on diagnostics from R. Reynolds	DRAFT	0.4	15th February 2003
C Donlon	Included comments and discussions following GHRSST-PP seminar at the Met Office, UK.	DRAFT	0.4.1	19th February, 2003
C Donlon	Added contributions from P. Minnett and B. Evans for validation and WP-ID7	DRAFT	0.5	21st February 2003
C Donlon	Included comments of F. Blanc, P. LeTraon based on MERSEA-IP discussions for data management and the GODAE data sharing project	DRAFT	0.5.1	23rd February 2003
C Donlon	Added discussions from the International GODAE Steering Committee meeting, EUMETSAT, 17-21st Feb 2003	DRAFT	0.5.2	26th February 2003
C Donlon	Added comments from M. Lynch (Curtin University, Perth, Australia) following 1 day discussion of GHRSST-PP at JRC on 28th Feb 2003	DRAFT	0.5.3	28th February, 2003
C Donlon	Added general review from P. LeBorgne comments	DRAFT	0.5.7	1 March 2003
C Donlon	Added additional P. LeBorgne comments following extensive discussions.	DRAFT	0.6	11th March 2003
C. Donlon	Included ideas and comments from various people at Met Office UK. (M. Bell (FOAM), R. Saunders (NWP satellite group)).	DRAFT	0.6.1	17th March 2003
C Donlon	Included comments from N. Rayner following meeting at Hedley Centre for Climate Change.	DRAFT	0.6.2	26th March 2003
C Donlon	Included comments following telephone discussion with G. Wick on merging procedures, grid size and general GDS structure.	DRAFT	0.61	28th March -2003
C Donlon	Discussions with C. Mutlow on data products, reference processor and services included.	DRAFT	0.7.1	29th March 2003
C Donlon	Added GHRSST-PP output grid specifications after extensive discussions of all Science Team	DRAFT	0.72	9th April 2003
C Donlon	Included collective comments and discussions following EGS/AGU/EGU Oral presentation, Nice.	DRAFT	0.7.3	12th April 2003
C Donlon	Included NGSSTv1.5 material from H. Kawamura	DRAFT	0.7.4	13th April 2003
C Donlon	Included discussion (at EGS Nice) with S. Pouliquen (IFREMER) for CORIOLIS and MEDS joint service support.	DRAFT	0.7.5	14th April 2003
C. Donlon	General review and edit	DRAFT	0.7.6	15th April 2003
C. Donlon	Included new specifications for Diurnal Confidence flag generation based on inputs and discussion with A. Stuart-	DRAFT	0.8.0	9th May 2003

	Menteth			
C. Donlon	Included comments from GHRSST-PP Science team. Substantial revisions to simplify the processing of data., Renamed to a processing specification. Dropped L4C and L4M data products from the specification. Revised the Functional breakdown diagrams and enhanced the section on MDB and validation. L2P data products now contain diurnal variation information	DRAFT	0.8.1	21st May 2003
C. Donlon	Included revisions based on input from D. May	DRAFT	0.8.2	28th May 2003
C. Donlon	Included output of the GHRSST-PP 4th workshop, Los Angeles USA, September – October 2003. Main changes: L2P and L4A format revised, deleted processing windows (PW), modified MPCF and IPCF, Revised L4A section, removed ERRLOG. Simplified entire document at the request of GHRSST-PP Science Team. Removed GRIB format specification. Added HR-DDS gridded format specification. Added new MDB data record definition tables. Completely revised the L4A section removing most details as requested by Science Team.	DRAFT	0.9	24th November 2003
C Donlon	Sent V0.99 for ST review and as input to Medspiration. Further revisions expected leading to the Medspiration SRR.	For ST Review	0.99	1st December 2003
C Donlon	Major revisions prior to Implementation of Medspiration project. Included substantial revisions to MW L2P section based on detailed discussions and comments from C. Gentemann at the CLIMAR-II conference and off line. Also included revised (and now correct) xml, FDGC DIF, metadata information (including (DTD specifications) for both the MMR and MDB based on comments from J. Piolle. netCDF CDL examples included. Included comments of I. Barton and G. Wick based on review of v 0.99. Updated the L4 section with comprehensive description of A. Stuart-Menteth diurnal variation parameterisation. Included a basic summary of the NCODA system provided by J. Cummings. Included general revisions and comments of P.LeBorgne. Generally tidied up for spelling and readability. Updated Annex 3 with Windsat information	Version 1.0 for use with ESA Medspiration project. This revision provides a reference baseline for the Medspiration project that will be upgraded in early February 2004.	1.0	22nd December 2003
J.F.Piollé & C Donlon	Signature of G Wick and general spelling typos, corrections to data management structures (MDB, MMR), comments form P. LeBorgne on prioritisation of sea ice and AOD information. Typos fixed, L2P and L4fnd product contents updated. A. Stuart-Menteth section updated.	V1.0	1.1	4th February 2004
C J Donlon	Revised coefficients for Stuart Menteth model. Update of JF-Piolle for CDL for L2P and L4A.	V1.0	1.2	9th February 2004
C Donlon, E. Armstrong, J. Vasquez, D. Hagan, Tim Liu, J.F.Piollé, G. Storkey	1 week intensive discussions and development of GDS and data management issues at JPL. Los Angeles, USA 9th – 13th February 2004 with J. Vasquez and E. Armstrong. Revised MMR content with E. Armstrong. Updated data provider tables for AIRS with D. Hagan. Updated Data provider tables for GOES with J. Vasquez and PO.DAAC. Updated all of the MDB file structure and content with E. Armstrong and JF-Piolle. Updated all of the MMR file structure and content with E. Armstrong and JF-Piolle. Revised L4SSTfnd content with E. Armstrong and J. Vasquez and JF-Piolle to be more useful based on experience at the JPL O.DAAC. Confirmed content with Tim Liu (JPL). Updated MetOffice data entry in Appendix 3 based on Dave Storkey input. Updated long_names of the HR-DDS sites	V1.0	1.3	20th February 2004
C Donlon, J.F.Piollé, P. LeBorgne, A Bingham (JPL).	Meeting at IFREMER Brest, 23 Feb 2004 to review and revise the GDS data management sections prior to the Medspiration SRR based on queries raised by the Medspiration SRR document (Release E). Revised the MMR and MDB structure slightly. Added date-time to L2 filename specification for files admitted to the GDS. New DTD and XML examples for all sections of MDB/MMR. Revised L2P format (lat/long definitions, time definitions, DT_min changed to DT_analysis, changed precision of SSI values, added L2_native_confidence_Value, added rules 2.2.2.8b and	V1.0	1.4	23rd February 2004

	2.2.1.5b, new definition of DT_min in section 2.2.1.4, added satellite_zenith_angle (for MDB consistency, added Rejection_Flag). Updated error log definitions and protocols in A7. Deleted section 2.5 and 2.6 as these are not required. Allow sub-sapling of 1km data sets prior toL2P data generation as required (for MODIS volume reduction). Refined definition of UHRSSTfnd to be regionally dependent as required. L4 file format revised to accommodate netCDF format (moved Skin_paramterisation_scheme to global variable, SSTskin are optionally computed based on RDAC/GDAC specific routines. Added L2P data file global attribute file_quality_index. Added information relevant to the delivery of data and metadata to the OCEANIDS system based on conversations with A. Bingham (JPL.DAAC).			
C J Donloon	Added NOAA/NESDIS aerosol optical thickness 100km weekly product as reference data stream R5	V1.0	1.4d	24th February 2004
C J Donlon	Reference document for Medspiration. Removed option to subsample or average input data to a lower spatial resolution. Integration time for SSI has been removed, but assumed that 3 or 6 hourly values will be used. L2P filenaming convension changed. Corrections to CDL example files. Corrections to Stuart-Menteth parametrisation equations. Added lower wind speed limit to Stuart-Menteth parameterisations. Corrections to the DT_analysis flow diagram Added rules 2.1.8 and 2.1.9 (land input and sea ice input rejection) Additions to HR-DDS sites from 4th GHRSST-PP workshop and Medspiration.	V1.0	1.5	5th March 2004
A Stuart-Menteth	Corrected DV parameterisation text. Added extra appendices for DV parameterisation to help implementation.	V1.0	1.5	19th March 2004
C Donlon	Corrections to equations 4.3.1.8, 4.3.1.10a and 4.3.1.10b	V1.0	1.5	23rd March 2004
C Donlon, J-F Piolle and E. Armstrong	Corrections to skin_time (variable name wrong, was time, should be skin_time) variable description of L4 data product CDL description. Final changes to MMR metadata tags. Updated details of AMSR-E sea ice concentration data from NSIDC. Removed references to DSR (should be DSD).	V1.0	1.5	24th March 2004
C Donlon	Additional corrections to Stuart-Menteth section	V1.0	1.5	26th March 2004
C Donlon	Additional corrections to Stuart-Menteth section	V1.0	1.5	2nd April 2004
C Donlon	Modifications to the MDB structure based on problems realised during implementation as part of the Medspiration project. Changes agreed during teleconference SOC/JPL/Met Office	V1.0		2nd December 2004
E Armstrong	Updated MMR_FR descripotion pg 118 tag should be Last_FR_revision_date for compatibility with MMR_DSD	v1.0		21st December 2004
J F Piolle, E Armstrong and C J Donlon	Revision of the MDB format and content (Appendix A4) based on initial experience within the Medspiration project and discussion within the MISST project,.	v1.0	1.6	11th January 2005
C J Donlon and E Armstrong	Revisions to MMR-DSD (Appendix A6) to extend character length definitions based on initial experience usin g the MMR_DSD.	v1.0	1.6	11th January 2005
C J Donlon and I J Barton	Revised table A5.4.1 to include GHRSST-Perth location (ghh069) and updated the coordinates of ghr061. Updated the MDB section with corrections	v1.0	1.6	17th February 2005
JF Piolle Robinson, W. Wimmer and D. Poulter	Updated the MDB section with corrections following discussions at MERSEA-IP WP2.2 meeting	v1.0	1.6	19th February 2005
E Armstrong	DSD description:- Dataset_Release_Date changed to ISO-8601 format and in FR example XML file FR_Creation_Date changerd to ISO-8601 format	v1.0	1.6	20th April 2005

Table of Contents

Executive Summary

The implementation of the Global Ocean Data Assimilation Experiment (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSST-PP) is based on a regional/global task sharing approach to the production and dissemination of GHRSST-PP sea surface temperature data products. A distributed data processing system is described that shares the operational data processing tasks that are necessary to generate and distribute common format high-resolution sea surface temperature (SST) data sets having global coverage in near real time.

A number of Regional Data Assembly Centres (RDAC) will generate regional coverage SST data streams in real-time by merging readily available satellite and in situ data products. Observational data products called L2P SST data products will be made available in near real time to GHRSST-PP users. Stringent quality control procedures are applied to input data and error statistics (bias and standard deviation) are assigned to each measurement. L2P data files will include estimates of the surface wind speed, aerosol optical depth, sea ice concentration and surface solar irradiance. These additional data fields are required to (a) interpret the quality of the near contemporaneous SST data themselves and (b) as input into advanced analysis systems. All available satellite SST data ingested by the GHRSST-PP will be formatted as netCDF L2P data products thereby providing a common self-describing data format. L2P data products have been developed for assimilation into ocean model systems and as

L2P data products are then assembled together at Global Data Analysis Centres (GDAC) where they are integrated and analysed to provide a global estimate of the SST at depth free of any diurnal variability (called the foundation temperature, SSTfnd) every 24 hours. Each SSTfnd data product will be accompanied by a suite of 2 hourly estimates of the SST in the top millimetre of the ocean providing information on the phase and magnitude of diurnal variation. GHRSST-PP users may access data products at either GDAC or RDAC in real time.

The GHRSST-PP Processing specification (GDS) is a common data processing specification that is implemented at each GHRSST-PP RDAC and GDAC. It represents a consensus opinion of the GHRSST-PP community on how to optimally combine satellite and in situ SST data streams within a globally distributed operational system. The GDS documents the minimum data processing that should be completed at each RDAC and GDAC in order that data products generated at different GDAC and RDAC can be used both individually and together with confidence. The GDS describes in detail the input and output data specifications, data processing procedures, algorithms and data product file formats **required** in order for the GHRSST-PP shared implementation model to function efficiently. It recognises that RDAC and GDAC must implement specific data processing procedures that account for data streams that are regionally specific (e.g., geostationary imagers). Furthermore, it is expected that RDAC may provide additional data products and services that satisfy regional user requirements (e.g., regionally specific analysed data products or ultra-high resolution data products). Such regional extensions to the GDS are only briefly considered in the GDS.

In order to verify that each processing centre has implemented the GDS correctly a reference GDS data processor will be used to process a set of data products using a GDS test data set. Each RDAC and GDAC should be able to provide the same output as the reference processor using the test data set as an input.

This document is the first version of the GDS (GDS v1.0) which is based on the First Report of the GHRSST-PP ISDI-TAG (Wick et al, 2002) and subsequent discussions at the Third GHRSST-PP Workshop, held at ESA/ESRIN, Frascati, Italy in December 2002 (Donlon et al., 2003b) and Fourth GHRSST-PP Workshop, Los Angeles, USA September 2003 (Donlon et al. 2004). This version of the GDS is focused on an initial data processing model with specific emphasis on implementing:

(a) An operational data exchange and delivery system between data providers, RDAC, GDAC and user communities by 2005,

(b) Definition and operational production of common format L2P SST data files for a number of near real time SST data products;
(c) A suite of regional L4 data products based on current best practice/knowledge analysis procedures that can be upgraded and refined based on experience. The analysis procedures should as far as possible capitalise on the synergy benefits of complementary measurements (accuracy, susceptibility to atmospheric clouds and aerosol loading, spatial and temporal resolution etc.)
(d) An initial global analysis system providing global coverage data products,
(e) An initial implementation of the GHRSST-PP High Resolution Diagnostic Data Set (HR-DDS) system and,
(f) A suite of initial test data products to be available to selected users at each RDAC and GDAC by mid 2004.

Much of the GDS is dedicated to issues of data exchange, management and operational considerations. It is a working reference document written primarily for the RDAC and GDAC teams as they physically realise, develop and, refine the GHRSST-PP demonstration system. As such, it may be considered a technical reference manual for the GHRSST-PP. The GDS will evolve throughout this process and a significant scientific upgrade of the processor is foreseen following the successful commission of the GDS at RDAC by regionally funded projects. The GHRSST-PP International Project office will work closely with all RDAC projects to refine and upgrade the GDS documentation and specifications as the GHRSST-PP project matures. Targeted upgrade paths are therefore clearly identified in a dedicated section of this document.

1. Introduction

The primary aim of the Global Ocean Data Assimilation Experiment (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSST-PP) is to develop and operate an operational demonstration system that will deliver high-resolution (better than 10 km and ~6 hours) global coverage SST data products for the diverse needs of GODAE and the wider scientific community. A new generation of SST data products will be derived and served to the user community by combining complementary satellite and in situ SST observations in near real time. A full description of the GHRSST-PP project is given in the GHRSST-PP Development and Implementation Plan (GDIP) which can be obtained from the GHRSST-PP project web server located at http://www.ghrsst-pp.org.

Figure 1.1 provides a simplified schematic overview of the GDIP. The GHRSST-PP is based on a distributed system in which the data processing operations that are necessary to operationally generate and distribute high resolution SST data sets having global coverage are shared by Regional Data Assembly Centres (RDAC). RDAC ingest, quality control and merge existing satellite and in situ SST data sources that are then used together to generate regional coverage quality controlled SST data products to the same specification (called L2P products), in real-time. RDAC data products are then assembled together at Global Data Analysis Centres (GDAC) where they are integrated and analysed to provide L4 global coverage data products.

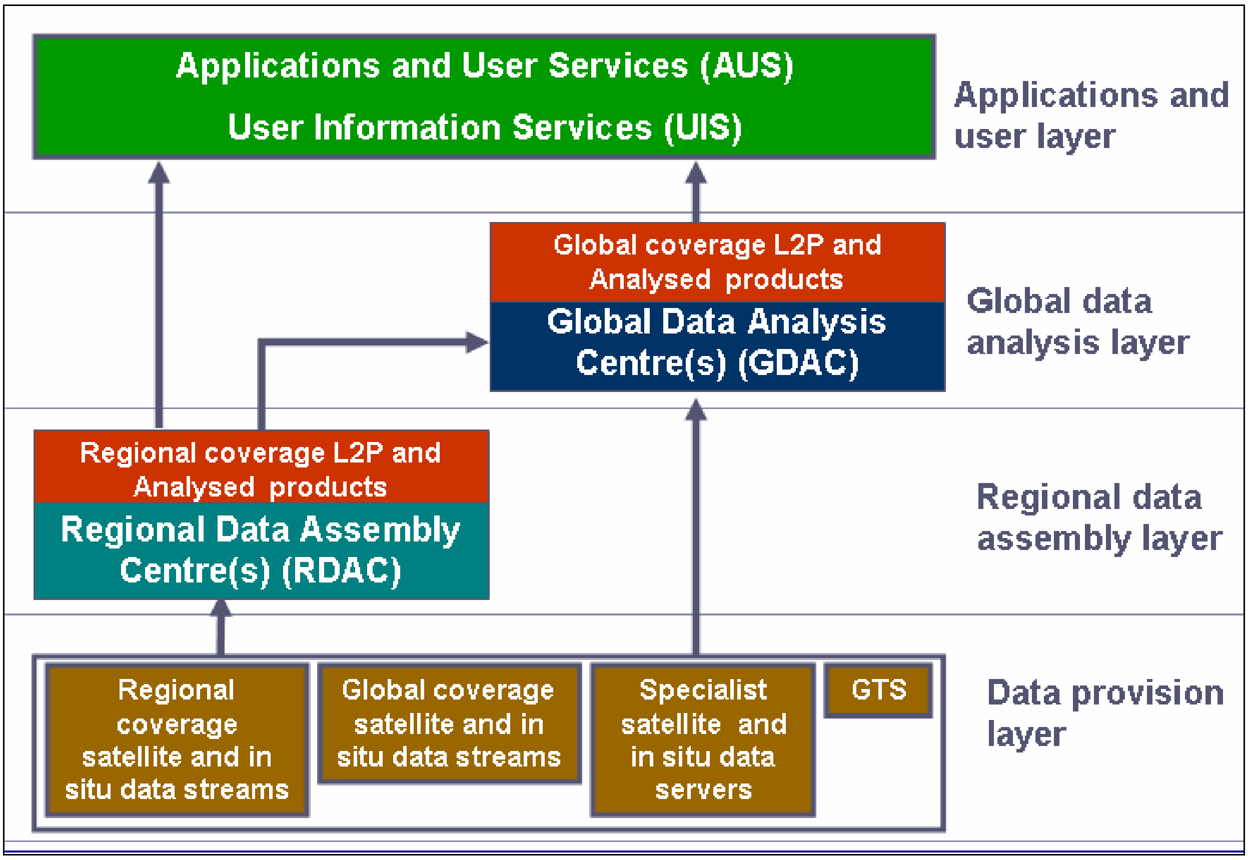


Figure 1.1. A schematic overview of the GHRSST-PP Development and Implementation Plan (GDIP) identifying the layered approach delivering data to applications. Red boxes indicate the output of GHRSST-PP data products.

Each RDAC will implement a processing system to generate and disseminate the GHRSST-PP data products that are common to all RDAC. Five GHRSST-PP RDAC projects are currently in preparation:

1. The New Generations SST project (NGSST) serving the western Pacific area replaced by a approximately defined by the Geostationary Meteorological Satellite (GMS, now replaced by GOES-9) footprint. The NGSST project is based in Japan.
2. The European Medspiration Project (MSP) serving the Atlantic area and European shelf seas.
3. The Ocean Forecasting Australia blueLink project serving the regional needs of the Australian region.
4. The Survey of the Environment Assisted by Satellite (SEASnet) program of the IDD serving the tropical oceans. The SEASnet project is based in France.
5. A project serving the SST needs of the USA under the US National Ocean Partnership Program (NOPP) under the general title of 'SST for GODAE'.

A GDAC centre is currently in preparation that will be implemented as a joint system by the Jet propulsion Laboratory Physical Oceanography Data Active Archive Center (PO.DAAC, http://podaac.jpl.nasa.gov) and the US-GODAE data server system at Monterey (http://www.usgodae.org). The PO.DAAC will be responsible for the data management of the GDAC including metadata, data serving and user interactions and the US-GODAE system will be responsible for the archive of GHRSST-PP data sets and the production of a global analysed SST fields initially using the US-Navy Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS). The US-GODAE server is dedicated to serving large data sets to operational ocean models. In addition, the extensive experience and capability of the PO.DAAC will ensure that GHRSST-PP scientific users are provided with excellent support, documentation and data access.

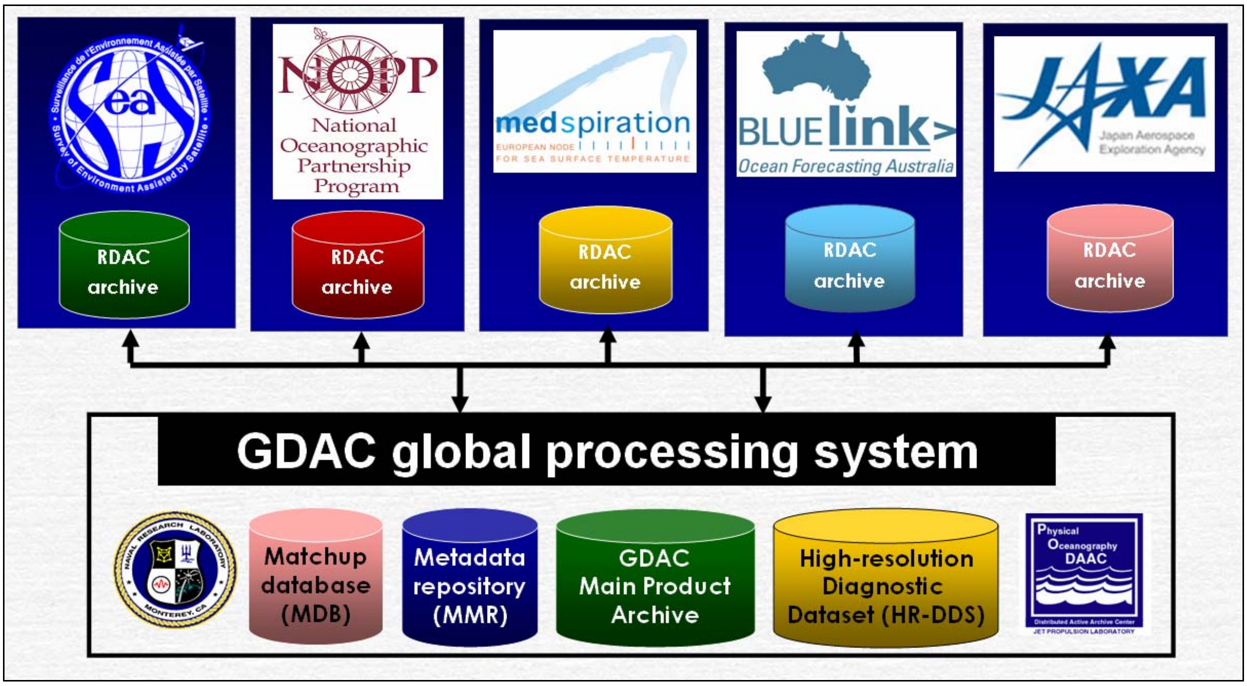


Figure 1.2 RDAC regional projects and their relationship to the GHRSST-PP GDAC jointly hosted by the PO.DAAC and US-GODAE.

In parallel, a second backup GDAC will be developed as part of the European Marine Environment and Security in the European Area (MERSEA) initiative (see http://www.ifremer.fr/merseaip/objectives.htm). The EU RDAC project Medspiration forms the initial pilot study for the MERSEA GDAC which will provide an operational backup processing system to the US GDAC described above. MERSEA aims to develop a European system for operational monitoring and forecasting on global and regional scales of the ocean physics, biogeochemistry and ecosystems. The prediction time scales of interest extend from days to months. This integrated system will be the Ocean component of the future European Global

Monitoring for Environment and Security (GMES) system. An initial configuration of the MERSEA EU GDAC is expected in 2006/7.

Figure 1.2 shows the configuration of the GHRSST-PP Regional/Global task sharing system that will implement the GDS. In addition to the global processor system, the GDAC also hosts the GDS (international) system-wide metadata repository (MMR), matchup database system (MDB), High Resolution Diagnostic Data Set (HR-DDS) and main product archive. RDAC will also maintain their own specific archives of L2P data products, HR-DDS, MDB and ultra-high resolution (~2km spatial grid) data products as required but the MMR will always reside at the GDAC.

1.1 The GHRSST-PP Data Processing Specification (GDS)

The GHRSST-PP Data Processing Specification (GDS) is a recommended **common** data processing specification that should be implemented at each GHRSST-PP RDAC and GDAC. It defines clearly the input and output data specifications, data processing procedures, algorithms and data product file formats that are used within the GDS and are thus **common to each GHRSST-PP RDAC and GDAC**. This is a prerequisite if the GHRSST-PP Global/Regional task sharing implementation framework is to function efficiently. For example, a common processing description is necessary to simplify documentation of data, facilitate exchange by sharing a common data format agreed by RDAC, GDAC and users, to avoid significant duplication of effort, to minimise reformatting of different data products derived by RDAC and to ease the integration of RDAC data to provide global coverage data sets at GDAC centres. Operationally produced data products will be improved by using additional data that are only available in a delayed mode together with extensive quality control procedures as part of the GHRSST-PP reanalysis (RAN) project (Casey et al., 2003).

GHRSST-PP RDAC centres must implement data processing procedures that account for specific aspects of regional coverage input data products (e.g., geostationary imagers). RDAC must also provide additional data products and services to satisfy regional user requirements (e.g., regionally specific analysed data products or ultra-high resolution data products) that will require research and development of new analysis, quality control, and data provision procedures. ***Regional extensions to the GDS are both necessary and critical to the preservation of regional identity and the proper evolution of the GDS***; without evolution, the data products and services provided by the GHRSST-PP are unlikely to satisfy developing user demands and concerns.

The first version of the GDS (v1.0) is focused on an initial data processing specification with specific emphasis on implementing:

(g) An operational data exchange and delivery system between data providers, RDAC, GDAC and user communities by 2005,
(h) Definition and operational production of common format L2P SST data files for a number of near real time SST data products;
(i) A suite of regional L4 data products based on current best practice/knowledge analysis procedures that can be upgraded and refined based on experience. The analysis procedures should as far as possible capitalise on the synergy benefits of complementary measurements (accuracy, susceptibility to atmospheric clouds and aerosol loading, spatial and temporal resolution etc.)
(j) An initial global analysis system providing global coverage data products,
(k) An initial implementation of the GHRSST-PP High Resolution Diagnostic Data Set (HR-DDS) system and,
(l) A suite of initial test data products to be available to selected users at each RDAC and GDAC by mid 2004.

It is not possible to specify exactly how the final version of the GDS will be implemented and operated without first having an initial 'Version-1' system in place to make an informed assessment. Neither is it desirable to have a GDS that is rigid without the ability to innovate based on experience.

The GDS is split into several work package (WP) sections that are supported by extensive technical Appendices. This format provides a framework that preserves the readability of the processing specification while providing extensive technical references that can be easily maintained and updates without affecting the overall structure of the GDS. Work packages develop a modular approach with clearly defined input and output parameters that greatly assist in the development of large multi-institute/national projects such as the GHRSST-PP. While the interface parameters for each work package will remain relatively static, considerable flexibility within a WP is maintained by this type of approach.

1.2 Document Scope

This document is the GDS version 1 which is based on the First Report of the GHRSST-PP In situ and Satellite Data Integration Technical Advisory Group (ISDI-TAG, Wick et al, 2002) and many subsequent discussions at the Third GHRSST-PP Workshop, held at ESA/ESRIN, Frascati, Italy in December 2002 (Donlon, 2003b) and further refined at the 4th GHRSST-PP Science Team Meeting, Los Angeles, USA in September 2003. It represents a consensus opinion of the GHRSST-PP community of how to pursue the optimal combination of satellite and in situ data streams within a globally distributed operational system to provide a new generation of global coverage SST data products. Much of the document is dedicated to issues of data exchange, management and operational considerations. For certain international groups (particularly the European Medspiration project), it provides a reference baseline that will be implemented and innovated by their project.

The GDS will evolve as each of the GHRSST-PP RDAC projects gain experience and knowledge throughout the Pilot Project and a significant scientific upgrade of the processor is foreseen following the successful commission of the v1.0 GDS. This is thus a working document written primarily for the RDAC and GDAC teams as they physically realise, develop and, refine the GHRSST-PP demonstration system. The document content will be constantly modified as initial approaches are tried and refined by the RDAC teams. But, in order to maintain consistency with funding agencies, published releases of the GDS will occur at regular intervals and represent reference baselines. These will carry an integer revision number (e.g., GDS-v1.0, GDS-v2.0 etc). As such, each may be considered a technical reference manual for the GHRSST-PP for a given time window based on the best current knowledge of the GHRSST-PP community, the Science Team, data providers and RDAC/GDAC projects. The GDS will be modified, refined and updated and published electronically on a regular basis by the GHRSST-PP International Project Office (GHRSST-PO).

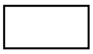
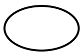
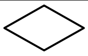
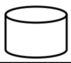
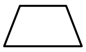
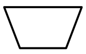
2. Notations, conventions and definitions used by the GDS.

The following sections describe the notations and conventions that are used throughout the GDS documentation. RDAC and GDAC implementation projects are expected to adhere to the nomenclature and style of the GDS in their own documentation as much as possible. This will greatly facilitate exchange of documentation between each centre.

2.1 Flow diagram symbols

The symbols described in Table 2.1.1 are used in all flow/functional breakdown diagrams.

Table 2.1.1 Symbols used in GDS flow and functional breakdown diagrams.

Symbol	Meaning
	denotes an algorithm step
	denotes an **algorithm step for which a further breakdown exists**
	denotes a **parameter**
	denotes an **interface parameter**
	denotes a **decision step**
	denotes a **data base**
	denotes the **start of a loop**
	denotes the **end of a loop**

2.2 Definition of data processing levels

The GDS uses the definitions provided in Table 2.2.1 when referring to data processing levels.

Table 2.2.1 Definition of satellite data processing levels.

Level	Abbreviation	Description
Level 0	L0	Unprocessed instrument and payload data at full resolution.
Level 1A	L1A	Reconstructed unprocessed instrument data at full resolution, time referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and geo-referencing parameters, computed and appended, but not applied, to the Level 0 data.
Level 1B	L1B	Level 1A data that have been processed to sensor units.
Level 2	L2	Geophysical variables derived from Level 1 source data at the same resolution and location as the Level 1 data i.e., satellite projection with geographic information.
Level 3	L3	Level 2 variables mapped on uniform space-time grid scales.
Level 4	L4	Results from analyses of lower level data (e.g., variables derived from multiple measurements). E.g., SST data sets generated from multiple source satellite data using optimal interpolation.

2.3 GDS data processing window specifications.

GDS data analysis processing activities are linked to Analysed Product Processing Window (APPW) periods within a 24 hour period that are defined in Table 2.3.1. A cutoff time (***Ω***) must be specified after which L2P data are no longer eligible for inclusion within the analysis. This must take account the time taken to process L2P data ($\Delta Time_{L2P}$), the time required to perform the analysis ($\Delta T_{analysis}$) itself and the Maximum desirable output time delay (T_{output}) using:

$$\Omega = T_{output} - \Delta T_{analysis} - \Delta T_{L2P} \qquad \text{(Eqn. 2.3.1)}$$

For example: if T_{output}= 12:00 at T+1, $\Delta T_{analysis}$=4h and, ΔT_{L2P}=2h then Ω= 06:00 at T+1. This means that L2P SST data for day T can be accepted until day T+1 at 06:00 UTC.

Table 2.3.1 Definition of GDS GDAC analysed product processing windows (APPW).

Name	Temporal coverage/Eligible input data	Time analysis is issued for	Latest actual output time
APPW	00:00 - 23:59 UTC Data cutoff: 06:00 (T+1)	12:00 UTC	As soon as possible following 23:59UTC but no later than 11:59(T+1)

2.4 GHRSST-PP definitions of sea surface temperature

Definitions of SST provide a necessary theoretical framework that can be used to understand the information content and relationships between measurements of SST made by different satellite and in situ instruments. The following SST definitions are defined and explained according to the consensus reached at the 2nd (Donlon, 2002b) and 3rd GHRSST-PP workshops (Donlon et al., 2003b). Each SST definition has been carefully considered by the GHRSST-PP Science Team in order to achieve the closest possible coincidence between what is ***defined*** and what ***can be measured operationally***, bearing in mind current scientific knowledge and understanding of how the near surface thermal structure of the ocean behaves in nature.

Figure 2.4.1 presents a schematic diagram that summarises the definition of SST in the upper 10m of the ocean and provides a framework to understand the differences between complementary SST measurements. It encapsulates the effects of dominant heat transport processes and time scales of variability associated with distinct vertical and volume regimes of the upper ocean water column (horizontal and temporal variability is implicitly assumed). Each of the definitions marked in the bottom right of the figure is explained in the following sub-sections.

2.4.1 The Interface SST (SSTint)

SSTint is a theoretical temperature at the precise air-sea interface. It represents the hypothetical temperature of the topmost layer of the ocean water and could be thought of as an even mix of water and air molecules. SSTint is of no practical use because it cannot be measured using current technology.

2.4.2 The Skin SST (SSTskin)

SSTskin is defined as the radiometric skin temperature measured by an infrared radiometer operating in the 10-12 µm spectral waveband. As such, it represents the actual temperature of the water at a depth of approximately 10-20 µm. This definition is chosen for consistency with the majority of infrared satellite and ship mounted radiometer measurements. SSTskin measurements are subject to a large potential diurnal cycle including cool skin layer effects (especially at night under clear skies and low wind speed conditions) and warm layer effects in the daytime (not shown in Figure 2.4.1).

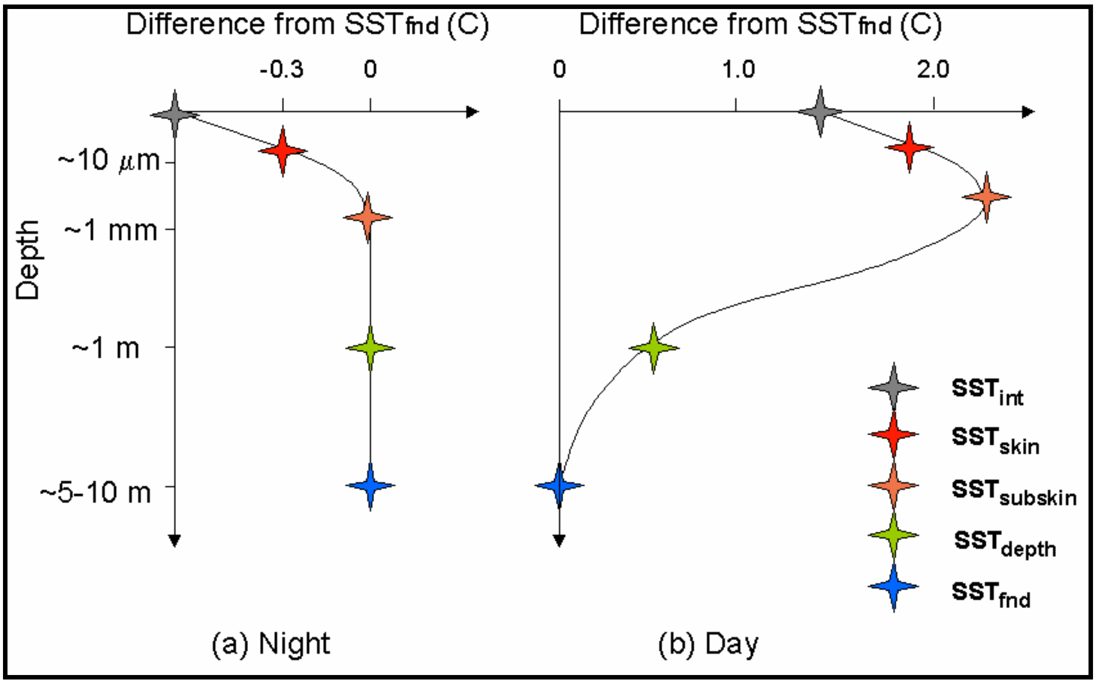


Figure 2.4.1 Schematic diagram showing (a) idealised night-time vertical temperature deviations from SSTfnd and (b) idealised day-time vertical temperature deviations from SSTfnd in the upper ocean.

2.4.3 The subskin SST (SSTsubskin)

SSTsubskin represents the temperature at the base of the thermal skin layer. The difference between SSTint and SSTsubskin is related to the net flux of heat through the thermal skin layer. For practical purposes, SSTsubskin can be well approximated to the measurement of surface temperature by a microwave radiometer operating in the 6-11 GHz frequency range, but the relationship is neither direct nor invariant to changing physical conditions or to the specific geometry of the microwave measurements.

2.4.4 The sea temperature at depth (SSTdepth)

All measurements of water temperature beneath the SSTsubskin are obtained from a wide variety of sensors such as drifting buoys having single temperature sensors attached to their hull, moored buoys that sometimes include deep thermistor chains at depths ranging from a few meters to a few thousand meters, thermosalinograph (TSG) systems aboard ships recording at a fixed depth while the vessel is underway, Conductivity Temperature and Depth (CTD) systems providing detailed vertical profiles of the thermohaline structure used during hydrographic surveys and to considerable depths of several thousand meters and various expendable bathythermograph systems (XBT). In all cases, these temperature observations are distinct from those obtained using remote sensing techniques and measurements at a given depth arguably should be referred to as 'sea temperature' (ST) qualified by a depth in meters rather than sea ***surface*** temperatures. The situation is complicated further when one considers ocean model outputs for which the SST may be the mean SST over a layer of the ocean several tens of meters thick.

SSTdepth or SST(z) is the terminology adopted by GHRSST-PP to represent an *in situ* measurement near the surface of the ocean that is typically reported simply as SST or "bulk" SST. For example SST_{6m} would refer to an SST measurement made at a depth of 6m. Without a clear statement of the precise depth at which the SST measurement was made, and the circumstances surrounding the measurement, such a sample lacks the information needed for comparison with, or validation of satellite-derived estimates of SST using other data sources. The terminology has been introduced to encourage the reporting of depth (z) along with the temperature.

2.4.5 The Foundation SST (SSTfnd)

The **foundation SST, SSTfnd**, is defined as the temperature of the water column free of diurnal temperature variability or equal to the SSTsubskin in the absence of any diurnal signal. It is named to indicate that it is the foundation temperature from which the growth of the diurnal thermocline develops each day. SSTfnd provides a connection with historical "bulk" SST measurements typically used as representative of the oceanic mixed layer temperature. This definition was adopted by GHRSST-PP at the Third GHRSST-PP Workshop (Donlon, 2003b) to provide a more precise, well-defined quantity than previous loosely defined "bulk" temperature quantities and consequently, a better representation of the mixed layer temperature. The SSTfnd product provides an SST that is free of any diurnal variations (daytime warming or nocturnal cooling). In general, SSTfnd will be similar to a night time minimum or pre-dawn value at depths of ~1-5 m, but some differences could exist. Only in situ contact thermometry is able to measure SSTfnd. SSTfnd cannot be directly measured using either microwave or infrared satellite instruments. Analysis procedures must be used to estimate the SSTfnd from radiometric measurements of SSTskin and SSTsubskin.

2.4.6 The diurnal cycle/variation of SST (DV)

For the GDS, a diurnal cycle refers to changes in vertical and horizontal distribution of SST throughout a 24 hour period and thus includes warm stratified layers ***and*** cool skin effects. Cool skin effects are typically more pronounced at night due to radiative cooling of the sea surface but may also occur during the day when the wind is light following a significant rainfall that may leave a cool freshwater layer on the surface of the ocean.

Warm layer effects are associated with environmental conditions characterised by low wind speed and strong insolation. A warm layer readily forms under such conditions effectively decoupling the surface layers (typically 0.1-3m deep) from the water beneath.

2.4.7 GDS operational output grid specifications

The following sections describe the output grid specifications for various GHRSST-PP data products.

2.4.7.1 Standard GHRSST-PP L4FND output grid specification

Table 2.4.7.1 describes the GDS preliminary output grid specification that is applicable to standard L4SSTfnd global coverage data products produced by the GDS.

Table 2.4.7.1 GHRSST-PP global coverage operational output grid specification

Description	**L4FND standard output grid specification**
Boundary	90S- 90N and 0°-360°W
Projection	Equidistant Cylindrical
Reference Ellipsoid	WGS84
Grid cell size	1/12° (9.28km at the equator)
Grid cell reference point	Centre of each grid cell. *e.g.,* for a grid cell at 35° referenced to the bottom left corner of the grid cell, the grid cell centre would be 35° + 1/24°

The GHRSST-PP Science Team adopted 1/12° as a definition for global coverage analysis data products to following the heritage of the AVHRR Pathfinder Ocean SST data products. It would have been equally valid to use a resolution of 0.1°.

2.4.7.2 Ultra-high Resolution GHRSST-PP L4FNDUHR output grid specification

Ultra-high resolution SSTfnd data products (L4FNDUHR) will be similar to L4FND data products but produced on a smaller grid specification described in Table 2.4.7.2. Note that RDAC are free to determine the exact resolution of these grids based on regional user requirements. A such, both metric and imperial units are provided as a guideline.

Table 2.4.7.2 GHRSST-PP regional coverage L4FNDUHR operational output grid specification

Description	L4FNDUHR standard output grid specification
Boundary	Regionally specified
Projection	Equidistant Cylindrical
Reference Ellipsoid	WGS84
Grid cell size	2km (~1/48° or ~0.02° at the equator)
Grid cell reference point	Centre of each grid cell. *e.g.*, for a grid cell at 35° referenced to the bottom left corner of the grid cell, the grid cell centre would be 35.01° (35° +1/96)

2.4.7.3 The GHRSST-PP High Resolution Diagnostic Data Set (HR-DDS) output grid specification

High Resolution Diagnostic Data Set (HR-DDS) will be constructed by remapping L2P data streams onto a uniform grid as specified in Table 2.4.7.3.

Table 2.4.7.3 GHRSST-PP HR-DDS operational output grid specification

Description	HR-DDS standard output grid specification
Boundary	Various defined in Table A5.4.1.
Projection	Equidistant Cylindrical
Reference Ellipsoid	WGS84
Grid cell size	0.01[1] (~1/96° at the equator)
Grid cell reference point	Centre of each grid cell. *e.g.*, for a grid cell at 35° referenced to the bottom left corner of the grid cell, the grid cell centre would be 35.005°

2.4.8 List of acronyms

Table 2.4.8 Acronyms and abbreviations applicable to the GDS

AATSR	Advanced Along Track Scanning Radiometer
AMSR-E	Advanced Microwave Scanning Radiometer - Earth Observing System
AMSR	Advanced Microwave Scanning Radiometer
AOD	Aerosol Optical Depth
APPW	Analysed Product Processing window
AUS	Applications and user services
Auxiliary data	Dynamic data that are used in the preparation of GHRSST-PP L2P data products including wind speed, surface solar irradiance, aerosol optical depth and sea ice.
AVHRR	Advanced Very High Resolution Radiometer
BT	Brightness temperature
Clw	Cloud liquid water content
CTD	Conductivity, temperature, depth (in situ ocean measurements)
DDS	Diagnostic data set
DODS	Distributed Oceanographic Data System

[1] The use of 0.01° as a the spatial resolution specification for HR-DDS data granules was introduced at the 4th GHRSST-PP Science Team meeting by the presentation of I. Barton in which example HR-DDS image data sets had been produced at a resolution of 0.01° latitude by longitude.

DRN	Data Ready Notification
DTD	Document Type Definition
DV	Diurnal Variation
ECMWF	European Centre for Medium-range Weather Forecasting
EOEP	Earth Observation Envelope Programme (ESA)
ERRLOG	Operational Error log
ESA	European Space Agency
ESL	Expert support laboratory (ESA)
EURDAC	European GHRSST-PP RDAC coverage
FOAM	Forecasting ocean assimilation model
GDAC	Global data analysis centre
GDIP	GHRSST-PP development and implementation plan
GDS	In situ and satellite data integration processing model
GHRSST-PO	International GHRSST-PP Project Office
GHRSST-PP	The GODAE High Resolution Sea Surface Temperature Pilot Project
GI	Global integration
GODAE	Global Ocean Data Assimilation Experiment
GOES	Geostationary operational environmental satellite
GOOS	Global ocean observing system
GTS	Global telemetry system
IODD	Input-output data definitions
IPCV	Infrared proximity confidence value
L2	Level-2
L2P	Level-2 data with added confidence flags after checking for gross errors, consistency and timelines. This family of data products provides the highest quality data obtained from a single sensor for a given processing window.
L3	Level-3
L4	Level-4
LAS	Live Access Server
MCSST	Multi-channel sea surface temperature
MDB	Match up database
MMR	Master Metadata Repository
MPCV	Microwave proximity confidence value
MSG	METEOSAT second Generation
NASDA	National Space Development Agency of Japan
NAAPS	Navy Aerosol Analysis Prediction System)
NCEP	National Center for Environmental Prediction (US)
NGSST	New Generation SST Project (Japan)
NOAA	National Ocean and Atmosphere Administration
NOPP	National Ocean Partnership Program
NWP	Numerical Weather Prediction
OPLOG	Operational data log
PEMSA	Production of Enhanced Multi-sensor SST Analysis Project (U.S.)
PO.DAAC	Physical Oceanography Data Active Archive Centre (U.S.)
PW	Processing Window
RA-2	Radar Altimeter-2
RAN	Re-analysis
RCN	Received Confirmation Notification
RD	Reference document (see section 1.5)
RDAC	Regional data assembly centre
Reference data	Pseudo static data and analysis products that are used by the GHRSST-PP (e.g., climatology maps, previous SST analysis (T-1))
RGTS	Regional and global task sharing

SEVIRI	Spinning Enhanced Visible and Infrared Imager
SQL	Structured Query Language
SSI	Surface Solar Irradiance
SSM/I	Special sensor microwave imager
SSES	Single Sensor Error Statistics
SST	Sea Surface Temperature
TAG	Technical advisory group
TMI	TRMM Microwave Imager
TRMM	Tropical Rainfall Mapping Mission
U	surface wind speed at 10m height
UHR	Ultra-high resolution
UIS	User information services
URD	User requirements document
VAL	Validation module of GHRSST-PP
WMO	World Meteorological Organisation
XBT	Expendable bathythermograph (In situ measurement of ocean temperature profiles)
XML	Extensible Mark-up Language

3. Overview of the GHRSST-PP Processing Specification v1.0 (GDS)

This section of the GDS provides a general overview of the processing model introducing how data will flow within the processor, the processor input and output data definitions (IODD), and the work packages (WP) that will be used to implement the processor. It may be used as a "map" of the GDS and identifies the interfaces between each specific work package. The work packages themselves are expanded in technical detail within subsequent sections and associated appendices of the GDS

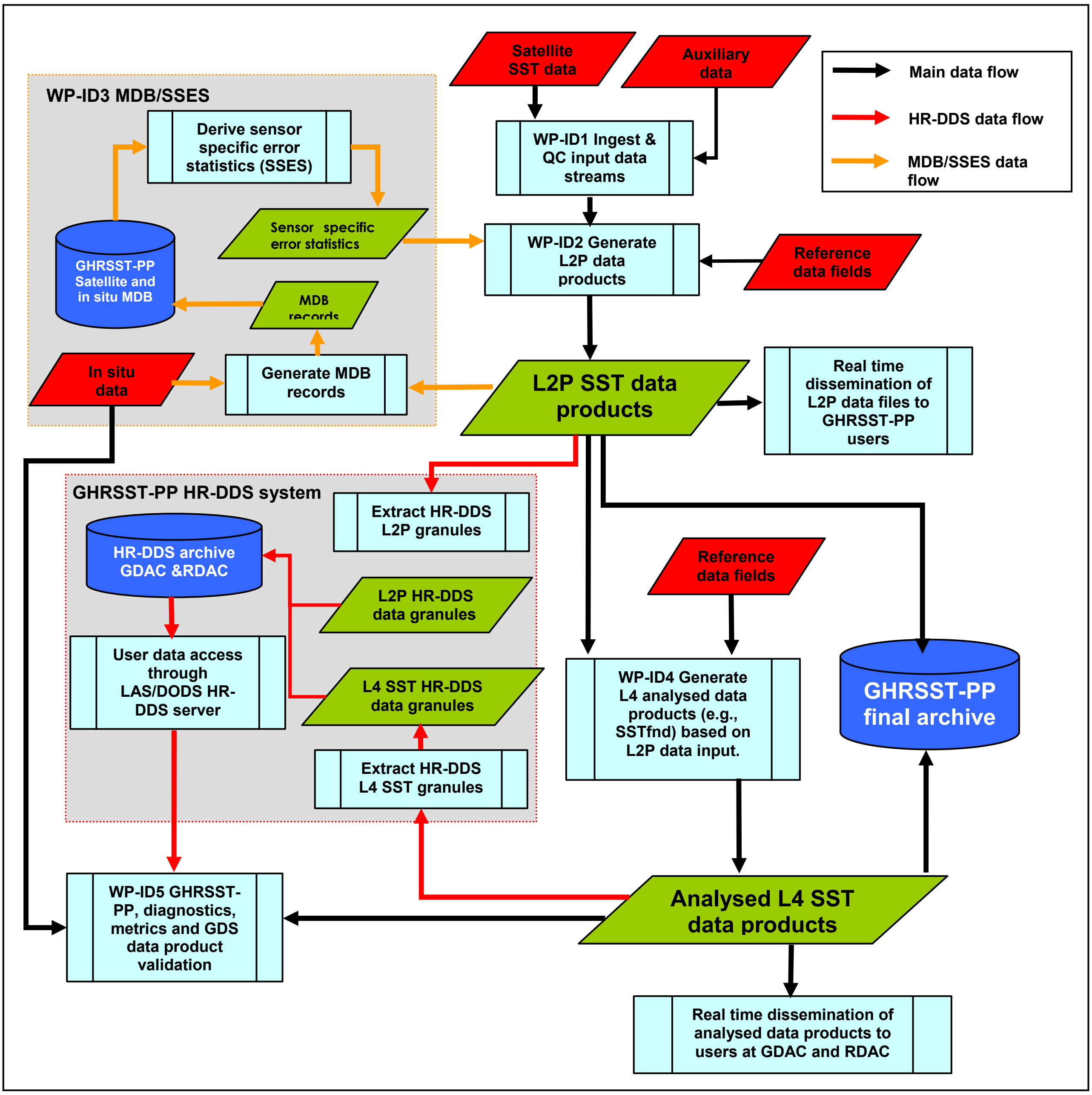


Figure 3.1.1 General overview of the GHRSST-PP GDS. Input data sources are shown in red, output parameters are shown in green and GDS processing steps with a further breakdown are shown in light blue. Database storage is shown in dark blue. For each output parameter a corresponding metadata record is generated and delivered to the GHRSST-PP Master Metadata Repository (MMR) system. The MMR system has been omitted to preserve clarity in this figure.

3.1 Summary description of the GDS

The GDS is designed to produce SST data products that satisfy the requirements of operational ocean forecast and prediction systems. Based on a user requirements survey (see the proceedings of the 3rd GHRSST-PP workshop and the Medspiration User Requirements Document for example, both available from the GHRSST-PO or http://www.ghrsst-pp.org), the main requirements are:

1. SST data are required operationally in near real time and should ideally be available within 6 hours of data acquisition at the satellite platform.
2. An error estimate (bias and standard deviation) and confidence data (such as an indication of atmospheric aerosol presence at the time of measurement) for each SST measurement including a bias and SD. is required by users and by the GHRSST-PP as a necessary precursor to data analysis.
3. A global ocean SST data product providing a best measure of the SSTfnd and SSTskin (including an estimate of diurnal magnitude and phase) is required at a minimum interval of 24 hours having a maximum spatial grid size of 1/12° (latitude x longitude).
4. Coastal modelling groups require ultra-high spatial resolution SSTfnd and SSTskin data products on a grid cell size of 1-2km and a timeliness of ~6 hourly.

Figure 3.1.1 shows a functional breakdown diagram of the GDS that identifies the flow of data, inputs and output channels and data. The GDS is split into work package (WP) sections to aid development and understanding of the system. The major work package components of the processing specification are shown in light blue together with their associated input and output data parameters that are shown in red and green respectively. The flow of data and information between each work package is indicated by the direction of connecting arrows.

The GDS data processing flow is designed to:

(a) Ingest and QC input SST data sets.
(b) Assign pixel confidence data, auxiliary data (wind speed, surface solar irradiance (SSI), aerosol optical depth (AOD), and sea ice) to each SST measurement.
(c) Assign error estimates (called Sensor Specific Error Statistics, SSES) to each SST measurement and format as a GHRSST-PP L2P output data file. L2P data sets should be available to the RDAC user community and GDAC processor ideally within 6 hours of data reception at the satellite sensor.
(d) Collocate L2P data records with in situ measurements and enter these data into the GHRSST-PP Match-up Database (MDB).
(e) Extract High Resolution Diagnostic Data Set (HR-DDS) data granules from L2P data files, reformat to HR-DDS data specifications and submit these data to the GHRSST-PP HR-DDS system.
(f) Consolidate L2P SST data sets every 24 hours and analyse all data within an Analysis Product Processing Window (APPW) to provide a ‘best estimate’ of the SSTfnd and SSTskin. Format analysed SST data sets as GHRSST-PP L4 data files and archive. L4 data products should be available to the user community no later than 12 hours following the end of an APPW. The analysis procedures, than must be developed, should aim to capitalise on the synergy benefits of complementary infrared and microwave SST measurements with particular emphasis on data combinations that lead to better accuracy, minimise the impacts of cloud/rain contamination and atmospheric aerosol loading, optimise spatial and temporal sampling resolution and address diurnal variability amongst others.
(g) Extract High Resolution Diagnostic Data Set (HR-DDS) data granules from L4 data, reformat to HR-DDS data specifications and submit these data to the GHRSST-PP HR-DDS system.
(h) For each data processing operation, create unique metadata records and submit these to the GHRSST-PP Master Metadata Repository (MMR).

(i) Archive L2P and L4 data in a GDAC archive for use in the GHRSST-PP Reanalysis Project (RAN).

Referring to Figure 3.1.1, RDAC first ingest, in real time, regional coverage satellite and in situ SST and auxiliary data streams from a variety of different data providers (WP-ID1 in Figure 3.1.1). Each input[2] satellite SST data stream is quality controlled to check for gross errors, consistency and timelines. Data that are unacceptable for further use in the GDS are rejected at this point. Each SST data file is then reformatted to a common GDS pre-processed data format (WP-ID2 in Figure 3.1.1) referred to as L2P and described in Appendix A1. This format is designed to allow easy sharing of data across the GHRSST-PP RDAC, provide a data product suitable for use by GHRSST-PP data analysis systems and ocean model data assimilation schemes, the conform to the GODAE data sharing project. L2P data products are the lowest-level, common format data products produced by the GHRSST-PP and provide the "building blocks" for all other higher level data products. Each L2P data product contains original SST measurements together with confidence fields and an error estimate (called single sensor error statistics (SSES)). L2P data products consist of the original SST measurements (that are not re-gridded or modified) together with additional confidence and error estimates. A separate L2P data product is produced for each sensor at each RDAC. Operational users that have requested this type of SST data product (see the GHRSST-PPO Development and Implementation Plan available at http://www.ghrsst-pp.org) may access L2P data products directly in real time from RDAC and GDAC. In addition, L2P data products provide a direct input to the GHRSST-PP Reanalysis project (Casey et al., 2003) and form the principal data archive for the GHRSST-PP.

A bias and standard deviation (sd.) error estimate is required for each [pixel] measurement before SST data originating from different sources can be properly assimilated into ocean model systems or analysed together to provide new SST data products. Assigning error estimates to SST measurements is a fundamental requirement of the GDS. Two types of error assignment are possible within a L2P data set:

1. Error estimates provided by a data provider as part of the L2/L4 input data
2. Statistical error estimates calculated from a match up database containing satellite and drifting/moored buoy in situ SST observations.

In the case of (2), there are clearly insufficient in situ observations to assign a unique error estimate for every satellite measurement. To make best use of the limited in situ data set that is available, in situ data spanning a particular time-window must be used to assign a statistical error estimate to all pixels within a data set. This is performed as a four stage process (WP-ID3 in Figure 3.1.1):

1. A quantitative "confidence value", having a value spanning (for example) 0 (no data, measurement is sea ice) to 5 (highest confidence in the SST measurement), is derived for each pixel measurement based primarily on an estimate of the surface wind speed, proximity to other cloud contaminated/rain contaminated pixels and, proximity to a SST reference climatology. Confidence values are optimized for likely errors from each sensor type (MW and IR).
2. L2P observations are then matched to near contemporaneous quality-controlled in situ SST observations in situ data and the resulting data record (together with any additional auxiliary data) are stored in a relational database system called the GHRSST-PP Match-up Database (MDB).
3. Error estimates for each SST measurement are then derived by analysing the MDB for a given time window (e.g., 1 week, 1 month, etc.). For each sensor and for each confidence value a Sensor Specific Error Statistic (SSES) is produced (i.e., for each sensor

[2] In the case of MW SST data sets, the use of level-2 (L2) as a descriptor of the data processing level is technically incorrect as these data are typically L3 gridded fields. Thus, in the context of L2P data descriptions, the input data are actually L2/L3 but in order to keep the terminology simple for L2P data files throughout the GDS the term L2P will be used for both L2 IR and L3 MW SST data sets.

and confidence value all data having the same confidence value are analysed to compute a mean bias and standard deviation).

4. The SSES bias and standard deviation computed in (3) for each confidence value then assigned to all [sensor specific] L2P data in the data set having a confidence value of the same value.

The format of data records held within the MDB is described in Appendix A4 and has been designed to be compatible with other SST MDB (*e.g.*, the Miami pathfinder MDB) in order to take advantage of these data and to further contribute to their development. Extensible Mark-up Language (XML) schemas have been developed to provide considerable flexibility in the content of MDB data records. MDB records will be ingested by a relational database system that will use sequential query language (SQL) scripts to analyse the database and produce SSES at periodic intervals triggered by user defined criteria (new data volumes).

GHRSST-PP L2P data products will serve the requirements of some operational ocean modelling groups in real time. However, many user applications and operational modelling groups request complete SST fields, free of gaps caused by clouds, rain or lack of data coverage, at regular intervals of between 6 and 24 hours. Most users require an estimate of the sub-surface SST field that is free of diurnal variability (the SSTfnd) that cannot be directly measured using current satellite techniques[3]. To address these needs, L2P data will used in an advanced analysis scheme (e.g., Reynolds and Smith, 1994; Reynolds et al, 2002; Guan and Kawamura, 2003, Murray et al. 2002, Murray et al. 1994, Fieguth et al., 1998; 2000, Menemenelis et al. 1997) that is designed to:

(a) Provide a daily global coverage combined SSTfnd product that builds on the synergy between complementary SST data streams having a grid cell size of 1/12° latitude x longitude (GDS-v1.0),
(b) Provide analysis system error estimates for each analysed grid cell,
(c) Account for differences in spatial and temporal sampling characteristics of each data stream,
(d) Account for gaps in coverage due to the presence of cloud, rain or lack of data,
(e) Account for SST diurnal variability and retrieve an estimate of the subsurface temperature field (SSTfnd),
(f) Provide a best estimate of the skin temperature of the ocean (SSTskin) for each grid cell,
(g) Provide a measure of diurnal variability (including phase and magnitude) within the data product time domain to accompany the SSTfnd estimate.

At 24 hour intervals, called Analysed Product Processing Windows (APPW) defined in Table 2.3.1, global coverage L2P data products are collated together at GDAC centres and used in an analysis procedure that will generate GHRSST-PP L4 analysed data products (WP-ID4 in Figure 3.1.1). L4 SST data products include the foundation SST (SSTfnd) including an estimate of SSTskin diurnal variability (magnitude and phase). Each L4 data product is formatted to a common GDS L4 data format (Appendix A1) that contains the analysis at each grid cell together with quality control and error statistic information. L4 data products will be available in real time to the GHRSST-PP user community no later than 12 hours after an APPW (see Table 2.3.1). MDB data records should be prepared for each L4 data product and sent to the GHRSST-PP MDB system for product validation activities (WP-ID5). Global analysis will take place at the GHRSST GDAC although regional analysis systems will be used at RDAC.

The GDS also prepares and delivers L2P and GHRSST-PP Analysed L4 data products for use within the GHRSST-PP high resolution diagnostic data set (HR-DDS), fully described in reference document GHRSST/14 "The HR-DDS implementation plan" and available at http://www.ghrsst-pp.org. Note that the GHRSST-PP HR-DDS is used for diagnostic analysis and validation of the GHRSST-PP GDS by the GHRSST-PP community and is completely separate from the GHRSST-PP

[3] Others require an estimate of the surface skin temperature that is more directly related to the overlying atmosphere and subject to considerable diurnal variability.

Match-up Database (MDB) system. The HR-DDS consists of L2P SST and auxiliary spatial data (e.g., model outputs, additional satellite data such as ocean colour or wind speed) that are extracted from data streams and re-gridded to a resolution of 0.01° x 0.01° latitude x longitude over small well defined geographical areas called HR-DDS granules (see Appendix A5). The HR-DDS provides a data resource that allows easy inter-comparison of separate data streams in an operational manner. It is implemented as a shared data resource where data access to will be through a Live Access Server (or variant) system.

The primary outputs of the HR-DDS are:

- Operational assessment of relative satellite SST bias changes (typically caused by an environmental event (e.g., volcanic eruption, atmospheric dust) or sensor problems).
- To test, validate and refine the methods and data products that are produced by the GDS itself.
- Provide a focus for commissioning the GDS and the production of on-going metrics required to assess the performance of the processor.

Each GDS work package has been designed to encapsulate a distinct suite of activities, common to all RDAC or GDAC that has a definite input and output definition (IODD). In this way, the exchange and use of all data products within the GDS is greatly simplified by referring to the work package interface IODD provided in section 3.2.

3.2 Input and output data definitions (IODD) for the GDS

Appendix A3 provides a complete description of the input data streams for the GDS and the main outputs of the GDS are listed in Table 3.2.1.

Table 3.2.1 Summary of GDS output

Product identifier	Descriptive name	Description	GDS section	Data format definition
L2P	L2 pre-processed data	Native satellite SST and auxiliary data that have been quality controlled and re-formatted to include confidence and error statistic data. <filename> is defined in Table A2.2	WP-ID2	Appendix A1.1
L4FND	GDAC L4 SSTfnd analysed data	Global coverage 1/12° spatial resolution SSTfnd analysed data products for each APPW	WP-ID4	Appendix A1.2
L4FNDUHR	RDAC L4 SSTfnd analysed data	Regional coverage SSTfnd analysed data products at ultra-high resolution (0.02° spatial resolution) for each APPW	WP-ID4	Appendix A1.2
SSES	Sensor specific error statistics	Mean bias and sd. Error statistical relationships to sensor specific confidence values.	WP-ID3	N/A
MDB	Match up data base record	Near contemporaneous satellite and in situ data match up record.	WP-ID3	Appendix A4
L2P HR-DDS granules	L2P HR-DDS data granules	A high resolution diagnostic data set (HR-DDS) granule (2° x 2° latitude x longitude area) extracted from a L2P data product and re-sampled to a 0.01 x 0.01 latitude x longitude grid spacing.	Appendix A5	Appendix A1.1
L4FND HR-DDS	L4SSTfnd high resolution DDS data granule	A high resolution diagnostic data set (HR-DDS) granule (2° x 2° latitude x longitude area) extracted from a L4FND data product and re-sampled to a 0.01 x 0.01 latitude x longitude grid spacing.	Appendix A5	Appendix A1.2
L4FNDUHR HR-DDS	Ultra-high resolution regional coverage L4SSTfnd high resolution DDS data granule	A high resolution diagnostic data set (HR-DDS) granule (2° x 2° latitude x longitude area) extracted from a L4FNDUHR data product and re-sampled to a 0.01 x 0.01 latitude x longitude grid spacing.	Appendix A5	Appendix A1.2
GDS System-wide (RDAC and GDAC) Data Set Description (MMR_DSD) metadata	Data Set Description Metadata (MMR_DSD) records for each Data set	Data Set Description (MMR_DSD) Metadata records for each data product used and generated by the GDS that form the parent for each MMR_FR archived at the GHRSST-PP Master Metadata Repository (MMR)	Appendix A6	Appendix A6.2

GDS System-wide (RDAC and GDAC) file Record (MMR_FR) metadata	File Record Metadata (MMR_FR) records for each L2P, auxiliary and L4 data set.	File Record (MMR_FR) Metadata records for each data product used and generated by the GDS that are sent and archived at the GHRSST-PP Master Metadata Repository (MMR)	Appendix A6	Appendix A6.2
GDS configuration files	System configuration files	Various system configuration files that are used to control data processing activities within the RDAC and GDAC GDS systems. A master copy of these is maintained at the GHRSST-PO.	Various	Specific to RDAC/GDAC
RDAC/GDAC error messages	Error messages that are relevant to the operation of the international GHRSST-PP project	Various RDAC and GDAC error messages generated by data processing systems. Some will be specific to GDAC/RDAC whereas others are generic in terms of content (e.g., RDAC is operational). These must be defined by RDAC and GDAC in close collaboration with the GHERSST-PO and the Science Team. Error messages should be of use to the entire GHRSST-PP community. Note that error messages specific to GDAC/RDAC systems are of limited value.	Appendix A10	Specific to RDAC/GDAC

WP-ID1 Ingestion of data streams

WP-ID1	Ingestion of Satellite and in situ data streams
Work Package number:	**WP-ID1**
Leader:	**GHRSST-PO and Data Management sub-group (Contact: Ed Armstrong, ed@seastar.jpl.nasa.gov)**
Aim	**To configure the GDS to successfully ingest near real time data streams from different data providers.**

1	Objectives: 1. Specify the data streams that will be considered by the GDS. 2. Specify a data provider responsible for each data stream. 3. Specify the primary and any secondary GDAD/RDAC entry points for each data stream. 4. Specify the limitations to the use of data arising from specific agreements pertaining to data use within the GHRSST-PP. 5. To specify the GHRSST-PP Master Metadata Repository Data Set Description (MMR_DSD) records for each data stream. 6. Specify the data transfer mechanism from data provider to GHRSST-PP primary entry point.
2	Description: This WP is dedicated to the ingestion of satellite and auxiliary data streams (including in situ and NWP fields) within the GHRSST-PP GDS. It describes how each data stream is documented within the GHRSST-PP data management system and provides guidance for RDAC/GDAC on the essential data ingestion/exchange/provision control mechanisms within the GDS. For each data stream, a MMR_DSD record will be generated and registered by the GHRSST-PO and data management sub-group to which all other L2P MMR File records (MMR_FR) that describe individual data files are registered in WP-ID2. If data are rejected by the GDS then an appropriate error message should be sent to the data provider stating why the data have been rejected. The error message should be copied to the ERRLOG (errlog@ghrsst-pp.org).
3	**Inputs:** 1. Data product definitions. 2. Data provider definitions and restrictions. 3. RDAC/GDAC data usage descriptions. 4. Data access and use agreements. 5. MMR_DSD records.
4	**Outputs:** 1. E-mail error message sent data provider explaining data ingestion/format problems. 2. E-mail to the GHRSST-PP ERRLOG explaining data ingestion/format problems. 3. MMR_DSD records for each data stream used by the GDS.
5	**Acceptance tests:** 1. Error message indicating data set rejection is correctly generated. 2. Error- message is correctly delivered to data provider.

	3. Error messages are correctly delivered to the GHRSST-PP ERRLOG. 4. MMR_DSD records prepared and registered correctly within the GHRSST-PP MMR system. 5. Data files are correctly renamed on acceptance into the GDS.

Figure 4.1 provides a functional breakdown diagram that identifies the main data processing tasks within WP-ID1. These are:

1. Establish real time operational access to data streams at RDAC and GDAC centres as appropriate for the processing centre (i.e. regional or global).
2. Create a Master Metadata Repository Data Set Description (MMR_DSD) record within the GHRSST-PP MMR for each data stream described in Appendix A3.
3. Evaluate each data files for integrity when ingested at an RDAC or GDAC facility.
4. Determine if the data file should be accepted into the GDS. If not, then raise an error report and send this to the data provider and to the GDS ERRLOG.
5. Rename the data file by prefixing the RDAC identification code to the filename.

Each sub task is described in the following sections.

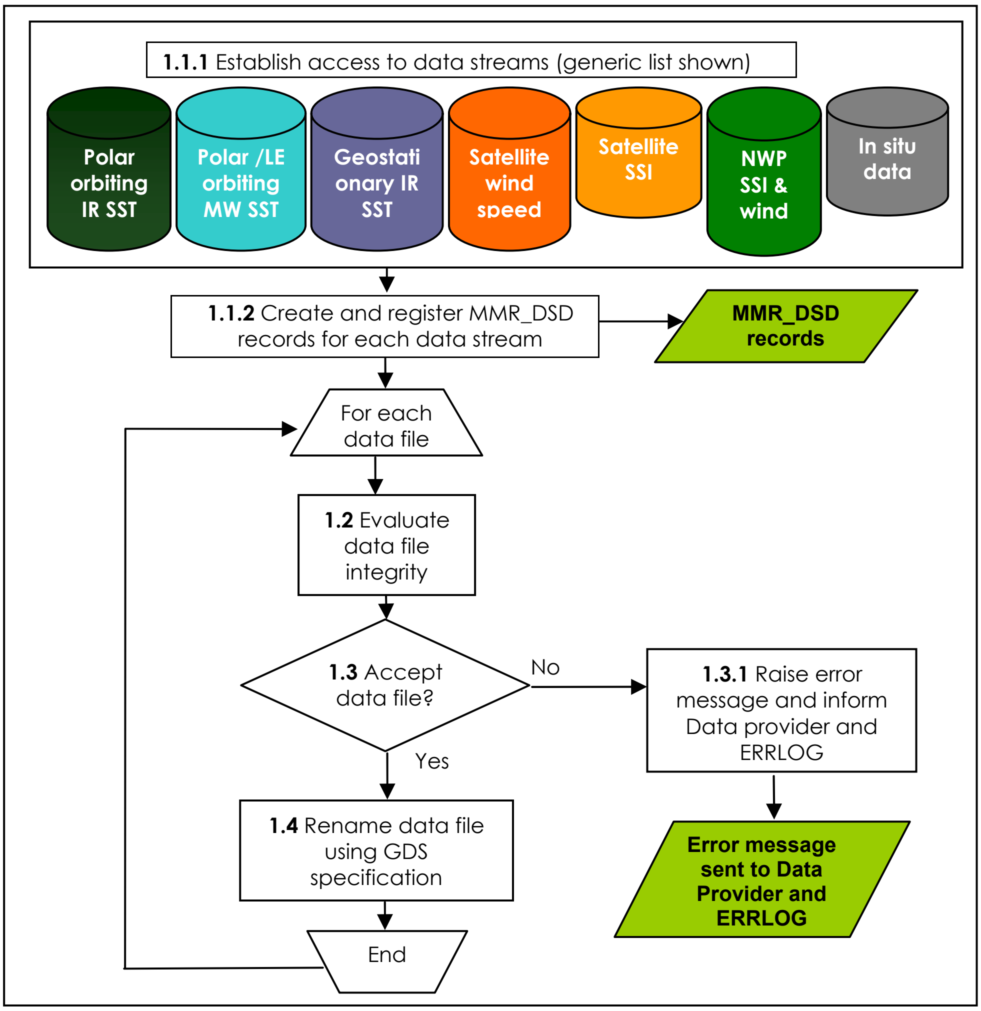


Figure 4.1 Functional breakdown of GDS WP-ID1 identifying each sub task.

WP-ID1.1.1 Access to input data streams

The GDS does not consider L0, L1a or L1B data streams (although L1B data streams may be available for certain HR-DDS sites if RDAC are able to supply these data). Satellite SST, auxiliary satellite data (including wind speed and surface solar irradiance (SSI)), in situ observations and NWP data are all required by the GDS to construct L2P data products and to produce L4 analysed SST data product outputs. Satellite data are ingested at RDAC and in some cases at GDAC centres. Each data stream will be formatted according to a well specified format provided by each data provider. Note that data streams from the same sensor may have different formats depending on the data provider. Table 4.1 provides an overview of the data sets that are available to the GHRSST-PP during the period 2003-2007 categorised by data type.

Table 4.1 Real Time global and regional satellite and in situ data streams available to the GHRSST-PP during the period 2003-2007. In situ data streams are those providing data within the upper 8 m of the water column. Acronym definitions are given in Appendix A2 Table A2.2.

Data type	Global coverage	Regional coverage
Operational[4] satellite data streams	AVHRR16-G (SST) # AVHRR17-G (SST) # SSM/I (wind) # QuickScat (wind) #	AVHRR16-L (SST) # AVHRR17-L (SST) # ATS_MET_2P$ (SST) SEVIRI# (SST, SSI) GOESE# (SST, SSI) GOESW# (SST, SSI) GOES-9# over Japanese region 0°N, 155°E (SST) MTSAT# (Launch in 2004)
Research satellite data streams	WindSat (SST, wind, not yet available) $ AMSRE (SST, wind) $ MODIS-A (SST, no NRT[5]) #	ATS_NR__2P# (SST) TMI$ (SST, wind)
Operational in situ data streams	None	ARGO VOSCL JMA DRI SOOP ODAS MAWS TAO PIRATA
Research in situ data streams	None	RADIOM CRUISE
NWP/Forecast data streams	ECMWF MetOffice NCEP NAAPS	

High data rate $ Low data rate

Appendix A3 describes in detail the data that will be used by the GDS. Appendix A3.1 defines the satellite SST data streams, Appendix A3.2 the auxiliary satellite data streams, Appendix A3.3 the in situ data streams and Appendix A3.4 the NWP/Forecast fields. A unique identification code for each data stream is given in Appendix A3 that will be referred to throughout the GDS. Also specified are the general characteristics of each data stream including the data provider,

[4] Operational in this context indicates that the data are available in near real time to the best ability of the data provider. It does infer that there are associated backup processing systems in place to ensure that data will be available 100% of the time.
[5] MODIS AQUA 1km SST data is now available for daytime orbits segment at the PO.DAAC. Night-time data may become available in the future.

data delivery mechanism, any data agreement that is applicable together with the primary and secondary RDAC/GDAC entry points. All of this information will be used by the GHRSST-PO and the GHRSST-PP Data management sub-group to generate MMR_DSD records for each data stream that is used by the GDS according to the specifications provided in Appendix A6 Table A6.2.1.

WP-ID1.1.2 Preparation of Master Metadata Repository (MMR) data set description records (MMR_DSD)

The data streams within the GDS are dynamic in the sense that new data sets will become available, certain data sets may take priority in given circumstances (e.g., aerosol contamination) and regional preferences based on the ease of data access will exist at RDAC. The GHRSST-PP Master Metadata Repository (MMR) is an online dynamic database that contains a metadata record for every data file that is produced by the GHRSST-PP. It provides a major source of information for RDAC and GDAC describing the availability of data files at any given instant in time. It also provides a catalogue for the GHRSST-PP archive.

The GHRSST-PP MMR system and the format of MMR_DSD and associated MMR File Records (MMR_FR), implementation of the MMR and delivery of MMR data records by RDAC and GDAC is described in Appendix A6. The GHRSST-PO in collaboration with individual data providers, RDAC and GDAC will prepare and maintain MMR_DSD records within the MMR system for each data stream used in the GDS. The information within the MMR_DSD must be kept current as these data will be used to monitor the availability of data, problems and issues that are relevant to operations and data quality.

WP-ID1.2 Evaluate the integrity of data files

When a data file enters a GDAC or RDAC processing facility following transfer from the data provider, it typically resides on a temporary staging area at an RDAC or GDAC centre. In some cases, the delivery may be incomplete; data connections could be lost, corruption of the data may have occurred during delivery, or the original file was incorrectly generated. ***The aim of the GDS at this stage is to identify a data delivery/data format problem and to initiate a dialog with the data provider to establish a solution as soon as possible.*** Individual RDAC and GDAC are responsible for implementing local procedures that are able to access data and a wide variety of tools will be used that are specific to each RDAC. RDAC/GDAC must also implement local procedures to assess the integrity of satellite data files, delivery protocols and reliability noting that these are likely to vary according to the data provider (i.e. the same data may be provided in a different format) and in time (data providers change file format, sensor is down, communication. problems). RDAC systems should focus on ***operational reliability and simplicity*** rather than elaborate procedures that are likely to introduce unnecessary complexity into the data processing system with minimal impact.

As a guide, the first GDS processing steps must evaluate the timeliness and integrity of each data file based on the following questions:

1. What is the status of the satellite system at present? Are there health warnings or environmental problems affecting the quality of data? This information should be present in the MMR_DSD record.
2. Was data transfer successful and is the data file complete? File sizes and CRC checks for example, will establish the situation.
3. Is the data file correctly formatted according to the specifications provided by the data provider? The data format could be randomly checked at the RDAC.
4. Was the delivery timely?

WP-ID1.3 Admit data file into the GDS system

Based on the assessment made by RDAC in WP-ID1.2 a decision is made to either accept a data stream or to reject the data stream from the GDS. The GDS specifies the following rules for this purpose:

- **Rule 1.3.1:** A data file should be admitted into the GDS processing system by an RDAC or GDAC only if the data file has been verified as complete, correctly formatted and that the delivery was timely with respect to the L2P delivery criteria of '***ideally***[6] available within 6 hours of acquisition at the satellite platform' and the analysed product processing window (APPW) defined in Table 2.3.1.

WP-ID1.3.1 Inform data provider and ERRLOG of data delivery/ingestion problem.

If a data file has been rejected by an RDAC/GDAC, an error must be raised and an e-mail message sent to the data provider that explains why the data file has been rejected from the GHRSST-PP GDS. These messages are used to:

(a) Establish a healthy feedback dialog with data providers
(b) To provide information to the GHRSST-PP that can be used to monitor GDS operations.

A separate error message should be sent for each data access attempt (i.e., if a second download is attempted and also fails two ERRLOG messages are generated). The data provider contact details and point of contact for error reports can be found in the L2P MMR_DSD records associated with each satellite data stream.

A copy of the error message should be sent via e-mail to the GHRSST-PP ERRLOG system using the address errlog@ghrsst-pp.org. Error messages are formatted as xml files according to the specification set out in Appendix A7.

WP-ID1.4 Rename data file to GHRSST-PP GDS specification.

In order to identify a data file that has been admitted into the GDS system by an RDAC or GDAC, the GDS recommends that each data file should be renamed by prefixing the processing centre code to the original filename according to the following filename specification:

`<processing centre code>-GDS-<original data filename>.<yyyymmddThhmmssZ>`

where **`<processing centre code>`** is defined in Appendix A2 Table A2.1 followed by an ISO format date-time string **`<yyyymmddThhmmssZ>`** for the time of data admission where

yyyy=year
mm=month of year (1-12)
dd=day of month (1-31)
T=identifier for time
hh=hour of day (0-23)
mm=minute of hour (0-59)
ss=second of minute (0-59)
Z=identified indicating time is expressed in UTC

[6] It may not be physically possible in some instances to obtain data within 6 hours of acquisition at the satellite platform due to data processing/delivery restrictions. Rather than exclude data due to poor timeliness, the GDS v1.0 cites the 6 hour timeliness criterion as an ideal target that may in some cases require the collaboration of data providers to attain.

Following the format described above, **`JAP-GDS-20010507tm.20040223T123211Z`** would refer to a REMSS TMI bmaps_03 data file[7] (20010507tm) that has been checked and admitted into the GDS at the Japanese RDAC (JAP) on the 23 February 2004 at 12:32:18 UTC.

Renaming a data file in this way provides a simple method to identify data that have passed the initial data delivery tests from those that have not while preserving the original data product filename. Although data files within this part of the GDS may have no visibility beyond the RDAC processor itself, the purpose of renaming files to a common specification is to facilitate error analysis and problem solving. If problems are found at a higher level within the GDS processing chain, accepted GDS files may be differentiated from other data files and traced back to RDAC/GDAC centres by simply noting the filename.

[7] In the case of Microwave gridded SST data files the use of L2 is maintained for readability although the data are in fact gridded L4.

WP-ID2: Generation of GHRSST-PP Level-2 Pre-processed Data files (L2P).

WP-ID2	Generation of GHRSST-PP Level-2 Pre-processed Data (L2P)
Work Package number :	**WP-ID2**
Leaders:	**GHRSST-PP Science Team and RDAC Project Managers (Contact: GHRSST-PO craig.donlon@jrc.it)**
Aim	**To process input data files and produce GHRSST-PP L2P SST data files in a timely manner.**
1	Objectives: 1. To extract confidence data provided with each input SST data file. 2. To quality control input SST data and derive GHRSST-PP pixel confidence data. 3. To determine if SST data are suitable for further use in the GDS. If not inform the data provider and generate an error message for the GHRSST-PP ERRLOG explaining why these data have not been used in the GDS. 4. Extract auxiliary data (wind speed, surface solar irradiance and aerosol optical depth) and collocate with SST data. 5. Create a GHRSST-PP netCDF L2Pdata file. Auxiliary fields will only be included prior to a cut-off time at which point the L2P data will be shipped 'as is'. 6. For each L2P data file, prepare and submit L2P MMR_FR to the GHRSST-PP MMR. 7. Submit appropriate error messages to the GHRSST-PP ERRLOG log (errlog@ghrsst-pp.org).
2	**Description:** This part of the GDS is dedicated to the production of L2P data files that consist of native SST observations together with additional pixel confidence data and auxiliary data stored in GDS netCDF format. L2P data products form the basic "building blocks" for all other GHRSST-PP data products and should ***ideally*** be made available to the GHRSST-PP operational user community in real time within 6 hours after the reception of data at the satellite sensor. Each input SST data file is assessed pixel by pixel, using gross error checking rules. A configuration file is used to store all QC parameter limits and thresholds that will be revised as required. Confidence data are extracted from native SST data streams if present and they are derived using auxiliary and reference data streams and sensor specific error statistics (SSES, generated by WP-ID7). For every L2P file that is generated, a L2P MMR_FR must be created and registered under the appropriate MMR_DSD at the GHRSST-PP MMR system as soon as possible. GHRSST-PP High Resolution Diagnostic Data Set (HR-DDS) granules are extracted from L2P data, formatted to a resolution of 0.01° latitude x longitude using a nearest neighbour re-sampling scheme and delivered to the GHRSST-PP HR-DDS system. If satellite SST data are unsuitable for further use in the GDS then a message is sent to the data provider and to the GHRSST-PP ERRLOG logs (errlog@ghrsst-pp.org) explaining why these data have not been used.

3	**Inputs:** 1. Infrared and Microwave satellite SST data streams. 2. Auxiliary satellite and NWP data streams. 3. Reference data streams. 4. GDS L2P configuration files. 5. Error checking rules.
4	**Outputs:** 1. GHRSST-PP L2P data files. 2. L2P MMR_FR metadata record. 3. Error message sent to data provider and to the GHRSST-PP ERRLOG log (errlog@ghrsst-pp.org) as required. 4. HR-DDS data granules extracted from L2P data.
5	**Acceptance tests:** 1. L2P output data products based on the GDS test input data set produced at each RDAC and GDAC should be identical to those produced by the GDS reference data processor. 2. L2P measurement data are identical to the input SST data but with additional error and QC). 3. HR-DDS data extraction is correct and data delivery interface is functional. 4. Gross error checking procedures successfully flag unsuitable data and admit useful data with minimal data loss. 5. Error messages to data provider and ERRLOG are correct and promptly delivered. 6. MMR_FR metadata records are submitted in a timely manner. 7. MMR_FR metadata records are accurate.
6	**Metric for performance assessment:** 1. L2P data files are correctly formatted. 2. L2P data are produced in a timely manner (ideally within 6 hours from reception at satellite platform). 3. L2P MMR_FR metadata records provided and ingested at MMR in real time. MMR rejection rate < 2%. 4. Data provider promptly notified in case of problems with data stream. 5. HR-DDS L2P extraction and format is correct. 6. ERRLOG messages are correct and only reject problematic data. 7. L2P data files are identically produced at each RDAC using the GHRSST-PP Test Data Set.

The purpose of this work package is to derive Level-2 Pre processed GHRSST-PP data files (L2P) through a process of quality control and data merging. The principles for setting the quality flags and pre-processing algorithms for assigning quality flags are outlined in this section. The L2P processor identifies and flags problematic satellite SST data at an early stage of the GDS processing chain leaving as much time as possible to correct faults and make use of the erroneous data later on in the processing system.

The main difference between the input SST data file and the output L2P data file is that additional confidence data and sensor specific error estimates for each pixel value have been included and the original SST data reformatted to netCDF format. Full orbit input data files may be split into ascending and descending files with a unique L2P output for each file.

L2P data products will be available to the operational user community in real time and ***ideally***, within 6 hours following reception at the satellite platform.

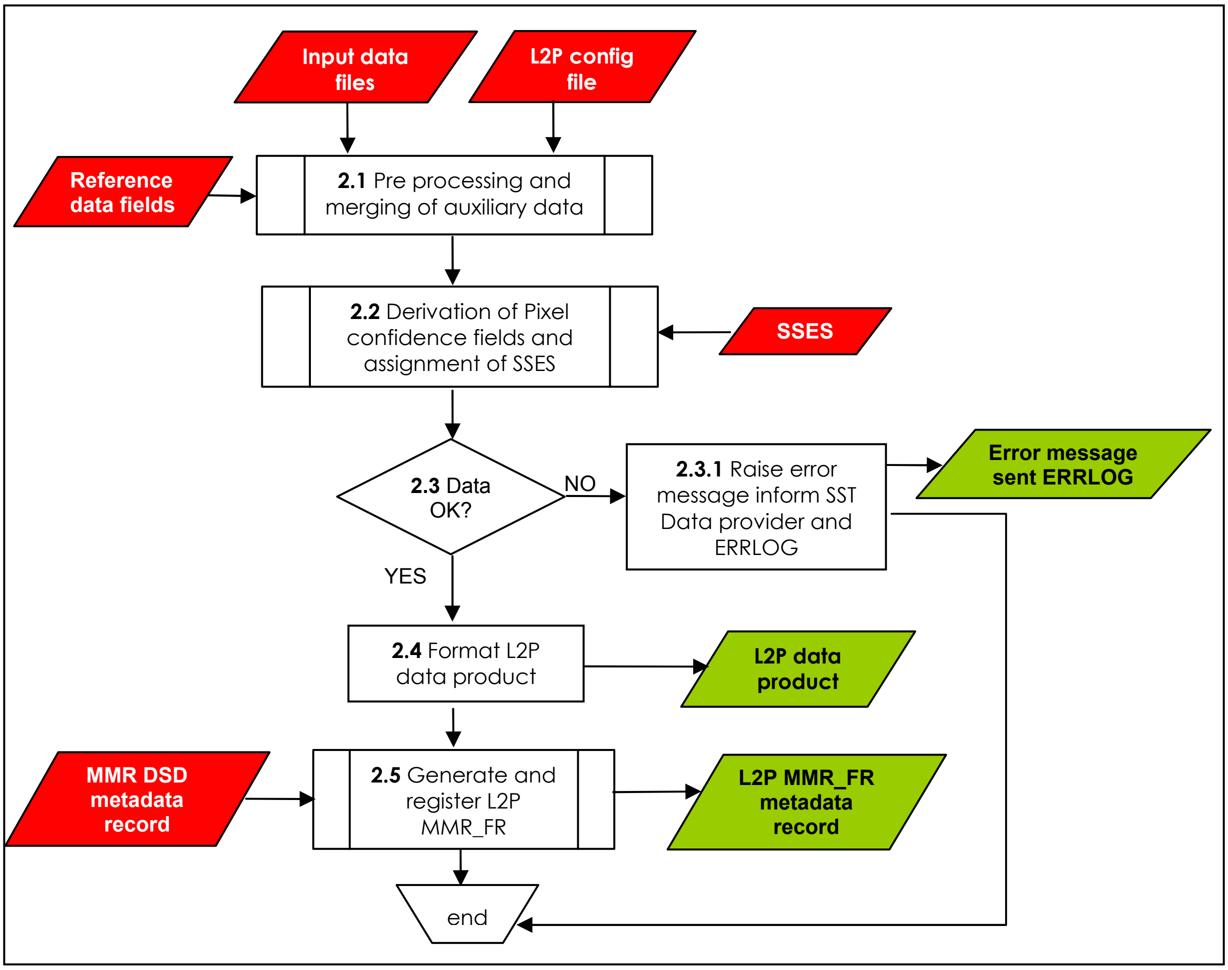


Figure 5.1 Functional breakdown of WP-ID2 identifying each major sub task.

Figure 5.1 shows a functional breakdown diagram for WP-ID2 (additional functional breakdown diagrams are provided for WP-ID2.1, WP-ID2.2 and WP-ID2.5 in the appropriate section below). In a series of pre-processing and quality control procedures, microwave and Infrared Satellite SST data are merged together with reference data and auxiliary data including satellite and NWP Wind speed, sea ice concentration, surface solar irradiance and aerosol optical depth estimates. If data processing problems are encountered the SST data file is rejected from the GDS, (e.g., values are outside of expected range, overwhelming cloud contamination) an error message is sent to the satellite SST data provider point of contact (as defined in the GHRSST-PP MMR_DSD record associated with the SST data file) explaining why the SST data have not been used in the GDS. A copy of the message is also sent to the GHRSST-PP error log system (errlog@ghrsst-pp.org).

If the SST data are accepted into the GDS, L2P additional pixel confidence data are then derived for every SST measurement. L2P confidence data are required to:

1. Derive a quantitative 'confidence value' for each SST measurement that is used to identify reduced quality data.
2. Derive and interpret of SSES values assigned to each SST measurement.
3. Force parameterisations that estimate diurnal variability.

Using confidence data, Sensor Specific Error Statistics (SSES) are assigned to each SST measurement based on analysis of the GHRSST-PP Match-up Database (MDB) system described in WP-ID3. A L2P data file is generated (format described in Appendix A1.4) which contains the original SST data together with pixel confidence data. L2P data files are formatted as GDS netCDF files (WP-ID1.2.3) that are fully described in Appendix A1.

For each L2P data file a MMR_FR metadata record (described in Appendix A6) is generated and registered at the MMR under the appropriate MMR_DSD. L2P data files and pixel confidence data are identical in format for all input SST data streams with differs between the IR and MW SST data exist. Using a single data format for all SST data files greatly facilitates SST inter-comparison and merging activities within the GHRSST-PP.

WP-ID2.1 Pre-processing and merging of auxiliary data.

This work-package describes in detail common pre-processing operations performed on each IR and MW SST data file. Figure 5.2.1 provides a functional breakdown diagram that identifies each common pre-processing pixel based operation. These are:

- Check that the SST value lies within acceptable limits, if not then reject the pixel value
- Compute the pixel measurement time
- Assign a fractional sea ice value if required
- Assign a wind speed value
- Assign a surface solar irradiance (SSI) value
- Assign an Atmospheric Optical Depth (AOD) value

The specific thresholds, limits and processing rules that are used to QC each data set are stored in a system wide L2P configuration file. This allows easy modification of critical values and other configurable parameters without the need for software code changes. As auxiliary data files are only used as ***indicators*** within the GDS, only simple QC tests are specified. It is left to the RDAC and GDAC to determine the level of QC that is required to use these data with confidence. Different pre-processing operations must be applied to IR and MW SST data and to each auxiliary data file.

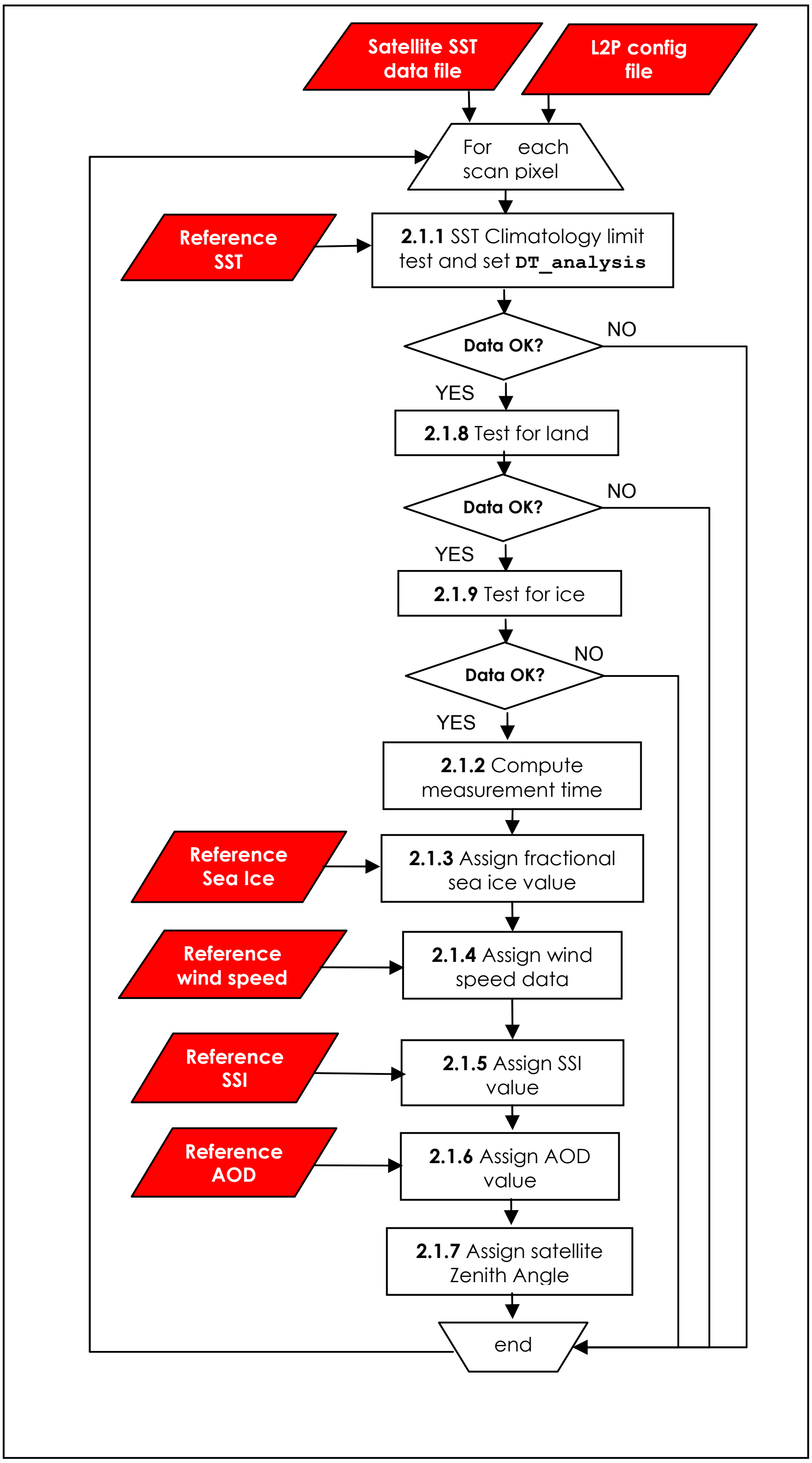


Figure 5.2.1 Functional breakdown diagram of pre-processing and merging operations common to IR and MW SST data files. A L2P configuration file is used to store appropriate thresholds and reference values.

WP-ID2.1.1 SST Climatology limit check

This test is used to set the L2P confidence data record variable **`DT_analysis`** defined in Table A1.2.2 in Appendix A1. It identifies data that fall outside of acceptable range of global SST values and identifies unrealistic SST values. The GDS specifies the following rules:

- **Rule 2.1.1a:** A valid SST observation must lie within a temperature range of **`LowSSTLimit`** < SST < **`HighSSTLimit`.** If the pixel SST value is out of these limits, the observation must be rejected from the final L2P data file. **`LowSSTLimit`** and **`HighSSTLimit`** are defined relative to the previous day SSTfnd analysis. The value of **`LowSSTLimit`** and **`HighSSTLimit`** are set in a L2P configuration file maintained at the RDAC and GHRSST-PP International Project Office as described in Appendix A9
- **Rule 2.1.1b:** **`DT_analysis`** is derived using:

$$\mathbf{DT\_analysis = SSTobs - SSTfnd_{(T-1)}} \qquad \textbf{(Eqn. 2.1.1)}$$

The **`DT_analysis`** value should be inserted into the **`DT_analysis`** field of the L2P confidence data record for the pixel in question as described in A1.2.3.5.

- **Rule 2.1.1c:** If a previous analysis SSTfnd data file is not available for use in Rule 2.1.1a and 2.1.1b, then the mean reference SST climatology (R9 in Table A3.3.1) should be used in its place.
- **Rule 2.1.1d:** If a measurement is rejected due to Rule 2.1.1a this should be recorded in bit 0 of the L2P **`rejection_flag`** variable as described in A1.2.3.18 so that a percentage rejection rate can be calculated for the input data file.

WP-ID2.1.2 Compute SST measurement time

Most SST data files provide the time at which a SST measurement was acquired but some do not. The GHRSST-PP GDS specifies that the time of measurement should be extracted or computed to the nearest minute and then coded as continuous UTC time coordinates using the following rules:

- **Rule 2.1.2a** The time of the earliest SST measurement within this data set should be coded as a continuous time coordinate specified as **seconds from 00:00:00 UTC January 1, 1981** (which is the definition of the **GHRSST-PP origin time**, chosen to approximate the start of useful AVHRR SST data record) and entered into the netCDF coordinate variable **`time`** (see A1.1.3).
- **Rule 2.1.2b** The time of SST measurement should be coded as the deviation in seconds from the netCDF coordinate variable **`time`** (see A1.1.3) and entered into the L2P record variable **`sst_dtime`**.

WP-ID2.1.3 Assign fractional sea ice value if required

Some SST data are contaminated in part or wholly by sea ice. This procedure is used to set the L2P confidence data record variable **`sea_ice_fraction`** defined in Table A1.2.2 in Appendix A1.

Some input SST data streams provide a flag to indicate that the SST measurement is contaminated by sea ice (e.g., AMSR-E). In cases where no such data are present in the input data stream a reference sea ice data product (defined in Appendix A3) should be used to assess the likelihood of sea ice contamination. The GDS specifies the following rules:

- **Rule 2.1.3a** If an input data set pixel fractional sea ice estimate exists, this should be used to set the **`sea_ice_fraction`** flag of the L2P confidence data record as described in A1.2.3.10.
- **Rule 2.1.3b** If an input data set pixel sea ice ***flag*** exists (i.e. indicating sea ice but not the fractional amount of coverage), this should be used to set the **`sea_ice_fraction`** flag of the L2P confidence data record to 100% according to the code values defined in Table A1.4.2 in Appendix A1. If the flag is a binary flag. a check should be made against the reference fractional sea Ice data set to determine if the pixel is likely to be 100% contaminated and the fraction reported by the reference data set should be set if it is < 100%.

- **Rule 2.1.3c** If an input data set pixel sea ice flag does not exist, and the pixel is located in or close to a region that may be ice contaminated, the reference sea ice data sets defined in Appendix A3 should be used to determine the value of the L2P confidence flag **`sea_ice_fraction`**.
- **Rule 2.1.3d** The L2P confidence bitfield **`sources_of_sea_ice_fraction`** must be set to indicate the data source used to set **`sea_ice_fraction`** using the values specified in Table 2.1.3 and described in A1.2.3.15
- **Rule 2.1.3e:** If the SST input data set includes a sea ice flag in the data stream, bit 4 of the L2P confidence data variable **`confidence_flag`** should be set for this pixel as described in A1.2.3.19.

Table 2.1.3 Code value for the source of the data used to assign the L2P confidence value `sources_of sea_ice_fraction`.

Code	sources_of sea_ice_fraction
0	No sea ice set
1	NSIDC SSM/I Cavialeri et al (1992)
2	AMSR-E
3	ECMWF
4	CMS (France) cloud mask used by Medspiration
…	Spare to be defined as required

WP-ID2.1.4 Assign wind speed data to L2P confidence record

Wind speed measurements are required within the GDS to derive L2P SST proximity confidence values, for use in diurnal variability correction schemes and to interpret the relationship between satellite and in situ SST data. At low wind speeds, especially in clear sky conditions, stronger diurnal variability is expected leading to higher surface layer temperature gradients and the potential for significant de-coupling of the skin/sub-skin SST from the SST at depth. Ideally a near contemporaneous wind speed measurement from satellite sensors should be used but this is impossible for all sensors due to the limited number of satellite wind speed sensors available. As a surrogate for a measured wind speed value, NWP estimates may be used as an indication of the surface wind speed.

The L2P confidence data record variable **`wind_speed`** defined in Table A1.2.2 contains a best estimate of the surface wind speed at the time of SST data acquisition. A near contemporaneous wind speed estimate[8] is assigned to this field as an indicator of the turbulent state of the air sea interface. The source of wind speed information must be specified in the L2P confidence data variable **`sources_of_wind_speed`** (defined in Table 2.1.4) and the difference in acquisition time between SST observation and wind speed measure must be specified in the L2P confidence variable **`wind_speed_dtime_from_sst`**.

The GDS specifies the following rules:

- **Rule 2.1.4a:** A 10m surface wind speed value should be assigned to each SST measurement pixel using the GDS L2P confidence data variable **`wind_speed`** described in Table A1.2.2 and described in A1.2.3.8. As a general rule for the GDS v1.0, only simultaneous wind speed measurements should be used otherwise NWP analyses should be used. Future revisions of the GDS will most likely use near contemporaneous wind speed observations from other satellite instruments.
- **Rule 2.1.4b:** The source of data used to set the L2P confidence data variable **`wind_speed`** should be indicated in the confidence variable **`sources_of_wind_speed`** as defined in Table 2.1.4 and described in A1.2.3.13.

[8] An NWP field may be used here.

- **Rule 2.1.4c:** In the case of microwave SST measurements, the use of a simultaneous microwave 10m wind speed measurements obtained from the same instrument providing the SST measurement may be used.
- **Rule 2.1.4d:** In the absence of a simultaneous surface wind speed measurement, an NWP estimated 10m surface wind speed (See Appendix A3, Table A3.1.1) should be used to set the L2P confidence data variable **`wind_speed`**.
- **Rule 2.1.4e:** The difference in time expressed in hours between the time of SST measurement and the time of wind speed measurement should be entered into the L2P confidence data variable **`wind_speed_dtime_from_sst`** as described in A1.2.3.9. In the case of an NWP field, this should be the central (mean) time of ant integrated value. A value of **25** should be used to indicate unknown time difference.

Table 2.1.4 Code value for the source of the data used to assign the L2P confidence value `sources_of_wind_speed`

Code	sources_of_wind_speed
0	No wind speed data available
1	AMSR-E data
2	TMI data
3	NWP: ECMWF
4	NWP: Met Office
5	NWP: NCEP
6	Reference climatology
9-15	Spare to be defined by RDAC as required

The following QC checks should be applied to all input wind speed data sets. The tests applied are simple tests designed to verify that wind speed data are within realistic limits. RDAC and GDAC may choose to implement more sophisticated QC procedures if deemed necessary.

WP-ID2.1.4.1 Gross wind speed limit check

This test identifies wind speed values that fall outside of acceptable range of global wind speed values. The GDS specifies the following rule:

- **Rule 2.1.4.1:** A valid wind speed observation must lie within a wind speed range of **`0 < wind_speed < HighWindSpeed`**. Wind speed values that fall outside this range are considered erroneous and should not be used within the GDS.
- **Rule 2.1.4.2:** If a measurement is rejected due to Rule 2.1.4.1 this should be recorded so that a percentage rejection rate can be calculated for the input data file.

The value **`HighWindSpeed`** is set in a L2P configuration file maintained at the RDAC and GHRSST-PP International Project Office as described in Appendix A9

WP-ID2.1.5 L2P Surface Solar Irradiance (SSI) value assignment

SSI measurements are required within the GDS to asses the magnitude and variability of significant diurnal SST variations, for use in diurnal variability correction schemes, as part of the L4 SST analysis procedures and to interpret the relationship between satellite and in situ SST data. Ideally a near contemporaneous SSI measurement from satellite sensors should be used but this is impossible for all areas due to the limited number of geostationary satellite sensors available. As a surrogate for a measured SSI value, NWP estimates may be used.

This test is used to set the L2P confidence data record variable **`surface_solar_irradiance`** defined in Table A1.2.2 in Appendix A1. A near contemporaneous SSI measurement, coded according to the definition provided in Appendix A1 Table A1.2.2 should be assigned to the L2P confidence data a record field **`surface_solar_irradiance`**. The source of SSI information must be specified in the L2P confidence data variable **`sources_of_ssi`** (defined in Table 2.1.5) and the difference in acquisition time between SST observation and SSI measure must be specified in the L2P confidence variable **`ssi_dtime_from_sst`.** The GDS specifies the following rules:

- **Rule 2.1.5a:** An integrated downwelling SSI measurement (e.g., derived from satellite measurements) should be assigned to each SST pixel value using the **`surface_solar_irradiance`** L2P confidence data variable. The SSI measurement nearest in space and time before the input pixel SST value should be used.
- **Rule 2.1.5b:** If no SSI measurement is available, an integrated SSI value derived from an NWP system (See Appendix A3) nearest in space and time to the SST measurement should be used to set the value of **`surface_solar_irradiance`**.
- **Rule 2.1.5c:** The source of data used to set the L2P confidence data variable **`surface_solar_irradiance`** should be indicated in the confidence variable **`sources_of_ssi`** as defined in Table 2.1.5.
- **Rule 2.1.5d:** The difference in time expressed in hours between the time of SST measurement and the time of SSI measurement/data validity should be entered into the L2P confidence data variable **`ssi_dtime_from_sst`**. The latest time of SSI measurement should be used for this field. A value of 25 should be used to indicate unknown time difference.

Table 2.1.5 Code value for the source of the data used to assign the L2P confidence value `sources_of_ssi`

Code	sources_of_ssi
0	No SSI data available
1	MSG_SEVIRI
2	GOES_E
3	GOES_W
4	ECMWF
5	NCEP
6	METOFFICE
…	For further use

The following QC checks should be applied to all input SSI data sets that are used at RDAC and GDAC. These tests are simple tests designed to verify that SSI data are within realistic limits. RDAC and GDAC may choose to implement more sophisticated QC procedures if deemed necessary.

WP-ID2.1.5.1 Gross SSI limit check

This test identifies data that fall outside of acceptable range of global SSI values. The GDS specifies the following rule:

- **Rule 2.1.5.1:** A valid integrated SSI observation/forecast must lie within a SSI range of **`0 < SSI < HighSSIValue`**. SSI values that fall outside this range should not be used within the GDS.
- **Rule 2.1.5.2:** If a measurement is rejected due to Rule 2.1.5.1 this should be recorded so that a percentage rejection rate can be calculated for the input data file.

The value **`HighSSIValue`** is set in a L2P configuration file maintained at the RDAC and GHRSST-PP International Project Office as described in Appendix A9

WP-ID2.1.6 L2P Aerosol Optical Depth (AOD) value assignment

This test is used to set the L2P confidence data record variable **`aerosol_optical_depth`** defined in Table A1.2.2 in Appendix A1. A near contemporaneous AOD measurement, coded according to the definition provided in Appendix A1 Table A1.2.2 should be assigned to the L2P confidence data a record field **`aerosol_optical_depth`**. The source of AOD information must be specified in the L2P confidence data variable **`sources_of_AOD`** (defined in Table 2.1.6) and the difference in acquisition time between SST observation and AOD measure must be specified in the L2P confidence variable **`AOD_dtime_from_sst`.** The GDS specifies the following rules:

- **Rule 2.1.6a:** A L2P AOD measurement (e.g., derived from satellite measurements) should be assigned to each SST pixel value using the **`aerosol_optical_depth`** L2P confidence data variable. The AOD measurement nearest in space and time to the input pixel SST value should be used.
- **Rule 2.1.6b:** If no AOD measurement is available, an AOD value derived from an NWP system or aerosol forecast (See Appendix A3) nearest in space and time to the SST measurement should be used to set the value of **`aerosol_optical_depth`**.
- **Rule 2.1.6c:** The source of data used to set the L2P confidence data variable **`aerosol_optical_depth`** should be indicated in the confidence variable **`sources_of_AOD`** as defined in Table 2.1.6.
- **Rule 2.1.6d:** The difference in time expressed in hours between the time of SST measurement and the time of AOD observations/analysis should be entered into the L2P confidence data variable **`AOD_dtime_from_sst`**. A value of **25** should be used to indicate unknown time difference.

Table 2.1.6 Code value for the source of the data used to assign the L2P confidence value `sources_of_AOD`

Code	sources_of_AOD
0	No AOD data available
1	NESDIS
2	NAVOCEANO
3	NAAPS
4-15	Spare (RDAC to suggest inputs here)

The following QC checks should be applied to all input AOD data sets. The tests applied are simple tests designed to verify that AOD data are within realistic limits. RDAC and GDAC may choose to implement more sophisticated QC procedures if deemed necessary.

WP-ID2.1.6.1 Gross AOD limit check

This test identifies AOD values that fall outside of acceptable range of global AOD values. The GDS specifies the following rule:

- **Rule 2.1.6.1:** A valid AOD observation/estimate must lie within a range of **`LowAODLimit < aerosol_optical_depth < HighAODLimit`**. AOD values that fall outside this range are considered erroneous and should not be used within the GDS.
- **Rule 2.1.6.2:** If a measurement is rejected due to Rule 2.1.4.1 this should be recorded so that a percentage rejection rate can be calculated for the input data file.

The value **`LowAODLimit`** and **`HighAODLimit`** are set in a L2P configuration file maintained at the RDAC and GHRSST-PP International Project Office as described in Appendix A9

WP-ID2.1.7 L2P satellite zenith angle value assignment

This test is used to record the **`satellite_zenith_angle`** variable for each measurement in a L2P data record. The GDS specifies the following rules:

- **Rule 2.1.7:** The satellite zenith angle for each input pixel measurement should be recorded in the L2P variable **`satellite_zenith_angle`** having a range of -90° to +90° with a precision of unit degrees as described in A1.2.3.17.

WP-ID2.1.8 Test for land input data

This test is used to record if an input pixel is land covered. The GDS specifies the following rules:

- **Rule 2.1.8:** If an input pixel is classified as land covered, the pixel should be rejected from the L2P data set and bit 6 of the L2P **`rejection_flag`** variable should be set.

WP-ID2.1.9 Test for sea ice input data

This test is used to record if an input pixel is ice covered. The GDS specifies the following rules:

- **Rule 2.1.9:** If an input pixel is classified as ice covered, the pixel should be rejected from the L2P data set and bit 4 of the L2P **`rejection_flag`** variable should be set.

WP-ID2.2 Derivation of L2P confidence data and assignment of Sensor Specific Error Statistics (SSES).

Due to characteristic differences between SST data sets derived from infrared and microwave sensors, the optimal assignment of error and confidence data for each measurement is different. In the case of infrared SST data sets, data are stratified according to their proximity to cloud and by wind speed (used to identify conditions favourable to stratification). Bias and standard deviation errors are assigned based on an analysis of the GHRSST-PP match-up database (MDB). A proximity confidence value is specified for each pixel based on the "proximity" of that pixel to several different criteria known to have a derogatory effect on the final SST value that is reported. The L2P confidence data record variable **`proximity_confidence`** defined in Table A1.4.2 in Appendix A1 is used to statistically derive and assign Sensor Specific Error Statistics (SSES) error estimate for each SST measurement. A long time series of data will be developed within the GHRSST-PP Match-up database (MDB) (containing limited geographical coverage of in situ SST and L2P data) that are analysed on a regular (weekly/monthly) basis according to sensor and **`proximity_confidence`** values to derive that provide a mean bias and standard deviation value that can then be applied to all other SST measurements having the same **`proximity_confidence`** value thereby maximising the use of the limited coverage of the in situ observations. The technique relies on the fact that SST measurements are generally of a reduced quality (and are therefore likely to have larger errors) the closer they are to other measurements that have been flagged as erroneous.

The GDS specifies separate schemes to derive **`proximity_confidence`** values for infrared and microwave SST data streams. For SST data derived from infrared radiometers, sub-pixel cloud and wind speed are the main factors influencing the quality of SST measurements. However, extremely large errors may be associated with atmospheric aerosol contamination due to dust storms or volcanic eruptions that have an immediate and widespread impact. This type of error is not considered in the GDS-v1.0 but will be included in a revised GDS once v1.0 is operating.

For microwave SST data streams, the proximity to rainfall, side lobe contamination in the proximity of land, high wind speed regimes modifying the surface emissivity of seawater in the 6-11 GHz frequency, radio frequency interference and proximity to sea ice are the main factors influencing the **`proximity_confidence`** value and the quality of SST retrieval. The approach adopted for the GDS-v1.0 is to flag suspect data in close proximity to rain, land, sea ice and at high wind speeds. Bias errors are assigned based on an analysis of the GHRSST-PP MDB. Standard deviation error increments associated with each error source are statistically derived based on look-up tables generated by the data provider (Remote Sensing Systems) and through analysis of the MDB.

WP-ID2.2.1 Derivation of L2P confidence data and assignment of SSES for Infrared SST data sets.

Figure 2.2.1.1 shows a functional breakdown diagram of the processing operations required to derive L2P confidence data and assign SSES for infrared SST data files. Initially, a series of quality control checks are performed on pixel-by-pixel bases that are designed to reject suspect data from the GDS. Proximity confidence values are then derived and SSES are assigned to each measurement. SSES are fully described in WP-ID3.

WP-ID2.2.1.1 Cosmetic data test

Some SST data streams contain pixel values that are cosmetic rather than measured in order to provide a complete data product[9]. The purpose of this test is to extract the status of an input SST flag (if such a flag is present in the native input data stream) and to reject all data that have been derived rather than measured by a sensor. The following GDS rule is specified:

- **Rule 2.2.1.1a:** If an input SST data stream includes information that indicates that a pixel value is cosmetic rather than measured, these data should be rejected from the L2P data file.
- **Rule 2.2.1.1b:** If a measurement is rejected due to Rule 2.2.1.1 this should be recorded in bit 1 of the L2P **`rejection_flag`** variable as described in A1.2.3.18 so that a percentage rejection rate can be calculated for the input data file.

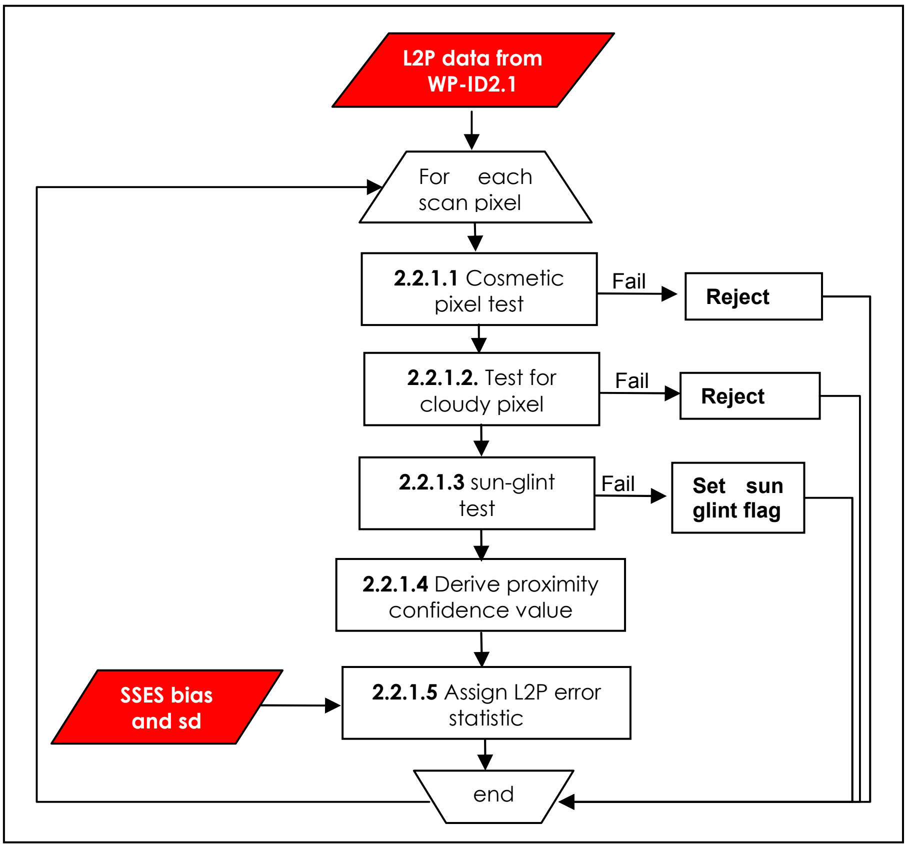


Figure 2.2.1.1 Functional breakdown diagram of data processing steps to produce L2P data records for infrared satellite SST data files.

WP-ID2.2.1.2 Cloud flag test

Many input SST data streams provide a flag or a data value indicating that the input pixel is cloudy. The GDS specifies the following rule:

- **Rule 2.2.1.2a:** If an input SST pixel data value or associated confidence data indicate the pixel to be cloudy, these data should be rejected from the L2P data file.

[9] For example, the ENVISAT ATS_NR__2P data stream may contain cosmetically filled pixels. These data are used to fill in data gaps across the swath following rectification of the curved ground swath made by the conical scan geometry used by the instrument.

- **Rule 2.2.1.2b:** If a measurement is rejected due to Rule 2.2.1.2 this should be recorded in bit 2 of the L2P **`rejection_flag`** variable as described in A1.2.3.18 so that a percentage rejection rate can be calculated for the input data file.

WP-ID2.2.1.3 Sun glint test

This test is used to set bit 5 of the L2P **`confidence_flag`** as described in A1.2.3.19. Depending on the local wind speed and sensor-sun view-geometry, sun-glint areas may be present in satellite SST data streams. Normal sun glint occurs in reflection from the oceans when the sun is behind and to the side of the satellite. Sun glint may degrade the SST value retrieved by affecting the quality of the cloud mask used by an infrared radiometer and consequently these regions should be flagged. Note that for polar orbiting satellite systems the problem is quite limited, but for geosynchronous orbit sensors it is unavoidable.

The GDS specifies the following rules:

- **Rule 2.2.1.3a:** If an indication of sun-glint is provided with an input SST pixel value, this should be used to set bit 5 of the L2P **`confidence_flag`** as described in A1.2.3.19.
- **Rule 2.2.1.3b:** If an indication of sun-glint is not provided with an input pixel value, an appropriate method should be used to set bit 5 of the L2P **`confidence_flag`** that is based on the geometry of sun and satellite position and surface roughness (e.g., Gardashov and Barla, 2001).

WP-ID2.2.1.4 Derivation of Infrared Proximity Confidence Values (IPCV)

The major source of error associated with infrared SST data sets is contamination by cloud due to poor cloud clearing techniques or particularly difficult environmental conditions (e.g., clouds are sub-pixel in size, SST is highly dynamic (coastal zone), thin cirrus cloud is present or low lying fog) that are difficult to detect using current algorithms. For the GDS v-1.0, IPCV are determined as a function of deviation from a minimum SST climatology (as cloud contaminated SST are cooler than surrounding waters) and distance from the nearest cloud flagged pixel (as smaller clouds are often found in close proximity to larger clouds). This type of approach is successfully used at the EUMETSAT O&SI SAF (Brisson et al., 2001) and as part of the US Navy NAVOCEANO operational SST processing system. The GDS-v1.0 IPCV scale is defined in Table 2.2.1.4.

Table 2.2.1.4 Infrared proximity confidence value (IPCV) scale definitions.

IPCF Value	Description
0	**Unprocessed:** Data that have not been classified (measurement indicates sea ice)
1	**Cloudy:**
2	**Bad:** Data that are probably contaminated by cloud.
3	**Suspect:** Data that may be contaminated by cloud or influenced by a cool skin effect.
4	**Acceptable:** Data that are reasonably distant from cloudy areas and in good agreement with the expected reference SST threshold.
5	**Excellent:** Data that are far from any cloudy areas and in good agreement with the expected reference SST threshold.
6	**Cool skin:** Data are far from any clouds but significantly cooler than the expected reference SST threshold. This could be isolated cloud, upwelling or other dynamical feature or a strong cool skin deviation.

The temperature difference **`DT_min`**, calculate between the pixel SST value (SSTobs) and the minimum SST climatology reference SST field[10], SSTref (defined in Appendix A3, Table A3.1.1 as R1), is derived using:

[10] This reference SST is SST1m derived as a 10 day "coldest" climatology from Pathfinder SST data products. SST values for a given location and season that are significantly beneath this climatological value indicate the presence of cloud contamination.

DT_min = SSTobs - SSTref **(Eqn. 2.2.1.4.1)**

Figure 2.2.1.4.1 presents a schematic lookup table describing how IPCV values are assigned. In this figure, four **`DT_min`** thresholds (**`IPCVThresh1 - IPCVThresh3`**) and two cloud proximity distances (**`IPCV_D1`** and **`IPCV_D2`**) delineate the set of data that correspond to the confidence scale provided in Table 2.2.4.1. It is expected that the IPCV scale will be expanded to identify an additional scheme that can be used to identify areas of diurnal stratification based on the difference between the previous analysis field and the observations and the SSI.

IPCV values are derived by considering the following GDS rules that have been found sufficient to identify most poor quality data (Brisson et al., 2001):

- **Rule 2.2.1.4a**: Typically, sub-pixel clouds often contaminate data immediately adjacent to cloudy areas. The spatial distance, **`D`** (expressed in km) of an SST observation to clearly flagged cloud contaminated areas is used to identify these measurements. Two threshold values set by the GDS, **`IPCV_D1`** and **`IPCV_D2`**, for each L2P data stream. The distance **`IPCV_D2`** is set on the assumption that cloudy pixels beyond this distance do not influence the SST measurement. The distance **`IPCV_D1`** is set assuming that cloudy pixels where **`IPCV_D1 < nearest_cloudy_pixel < IPCV_D2`** have a reasonable probability of being cloud contaminated. Pixels within a distance less than **`D1`** are assumed to have a high probability of cloud contamination. **`D`** values will depend on the spatial resolution and the geometry of each individual satellite sensor and must be established for each sensor. For example, **`D`** values for geostationary observations of 5km will be different from polar orbiting radiometer data of 1.1km.

Threshold values for `DT_min`, (K)>				
> `IPCVThres1`	**3 (Suspect)**	**3 (Suspect)**	**5 (Excellent)**	
< `IPCVThres1`	**2 (Bad)**	**2 (Bad)**	**4 (Acceptable)**	
> `IPCVThres2`	**1 (Cloudy)**	**2 (Bad)**	**4 (Acceptable)**	
< `IPCVThres3`	**1 (Cloudy)**	**(Cloudy)**	**6 (Suspect, Cool skin, upwelling, riverine inputs etc.)**	
	`IPCV_D1` (close)	**`IPCV_D2` (near)**	**> `IPCV_D2` (far)**	**Distance from nearest cloudy pixel (km)**

Figure 2.2.1.4.1 Schematic diagram showing how IPCV values are derived based on the deviation from a minimum SST climatology defining the coolest expected SST value and the distance from the nearest cloud flagged pixels.

- **Rule 2.2.1.4b:** The second IPCV test uses **`DT_min`**, the deviation from a minimum SST climatological value defined by equation 2.2.1.4.1. Three threshold values for **`DT_min`** will be set by the GDS L2P configuration file **`IPCVThresh1 - IPCVThresh3`**.
- **Rule 2.2.1.4c:** A static map of coastal regions typically influenced by cold waters (e.g., upwelling, river inputs) may be used to decide if a pixel is cloud contaminated in the case of IPCV value 6. Maps will be developed by RDAC as appropriate for specific regions of interest. As an initial minimum guideline for use with L2P data products, upwelling maps should be derived at 10 km resolution or better. For use with UHR data products, upwelling maps should be defined at 2 km or better.

Note that the actual values for **`IPCVThres1-3`** and **`IPCV_D1`** and **`IPCV_D2`** are expected to vary for each sensor (Brisson et al., 2001) and the exact value of these parameters is a subject of on-going research and development within the GHRSST-PP. **`IPCVThres1-3, IPCV_D1`** and **`IPCV_D2`** values will be maintained by the GHRSST-PP International Project Office and stored in the L2P configuration file that is specific to each RDAC or GDAC as described in Appendix A9. The derivation of these values will initially use best estimates until a sufficiently large population of data within the GHRSST-PP MDB is available (e.g., those suggested by Brisson et al. (2001). The filename of this configuration file should be written into each L2P file using the optional global netCDF attribute "References" as described in Appendix A1.2.1. It is expected that the methodology used to derive IPCV may be modified in the GDSv2.

WP-ID2.2.1.5 Assign Sensor Specific Error Statistic
This processing step is used to assign the L2P confidence data record variable **`sses_bias_error`** and **`sses_standard_deviation_error`** defined in Table A1.2.2 in Appendix A1. The GDS specifies the following rules:

- **Rule 2.2.1.5a:** If pixel bias error and sd error statistics are provided by a data provider, these values may be assigned to the L2P **`sses_bias_error`** and **`sses_standard_deviation_error`** variables respectively as described in A1.2.3.3 and A1.2.3.4. In this case, bit 6 of the L2P **`confidence_flag`** variable should be set as described in A1.2.3.19 because these error estimates are independent of GHRSST-PP methods.
- **Rule 2.2.1.5b**: If bit 6 of the L2P **`confidence_flag`** variable has been set, it is strongly advised that the associated **`proximity_confidence`** value is also updated as described in A1.2.3.20 by the data provider. The values should be fully compliant with the GDS IPCV scale. In addition, bit 7 of the L2P **`confidence_flag`** should be set as described in A1.2.3.19 to avoid inconsistency.

If native **`sses_bias_error`** and **`sses_standard_deviation_error`** variables values are not available or an RDAC chooses to ignore the error statistics provided by a data provider, these must be assigned based on the L2P **`proximity_confidence`** variable using sensor specific error statistics (SSES) as described in WP-ID3. The GDS specifies the following rule:

- **Rule 2.2.1.5c:** If pixel **`sses_bias_error`** and **`sses_standard_deviation_error`** error statistics are not available for SST measurements, these values should be assigned the most recent SSES bias and sd values for the each data stream and L2P confidence data record **`proximity_confidence`** value. In addition, bit 6 and bit 7 of the L2P **`confidence_flag`** should be set to zero as described in A1.2.3.19.

It is expected that the IPCV and SSES schemes will develop considerably throughout the GHRSST-PP.

WP-ID2.2.2 Quality control of satellite MW SST data sets and assignment of pixel confidence data.

SSES and L2P confidence and data must be derived for every MW SST measurement contained within an input SST data stream. The processing is different from infrared data streams since the native MW fields already contain wind speed, rain, and other ancillary parameters required to calculate the coded bitfields identifying suspect pixels. Specific thresholds, limits, reference files, and processing rules that are used to process each data set are stored in a system wide L2P microwave L2P configuration file. This allows easy modification of critical values and other configurable parameters without the need for software code changes.

Figure 2.2.2.1 shows a functional breakdown diagram of the processing steps required to derive confidence data and assign SSES to microwave SST measurements. The output of these processing steps is a GHRSST-PP L2P netCDF data file described in Appendix A1.4. Table A1.4.2 in Appendix A1 describes the format of GDS L2P pixel confidence data records. The following sections describe in detail the MW QC tests and pre-processing operations performed on each data file.

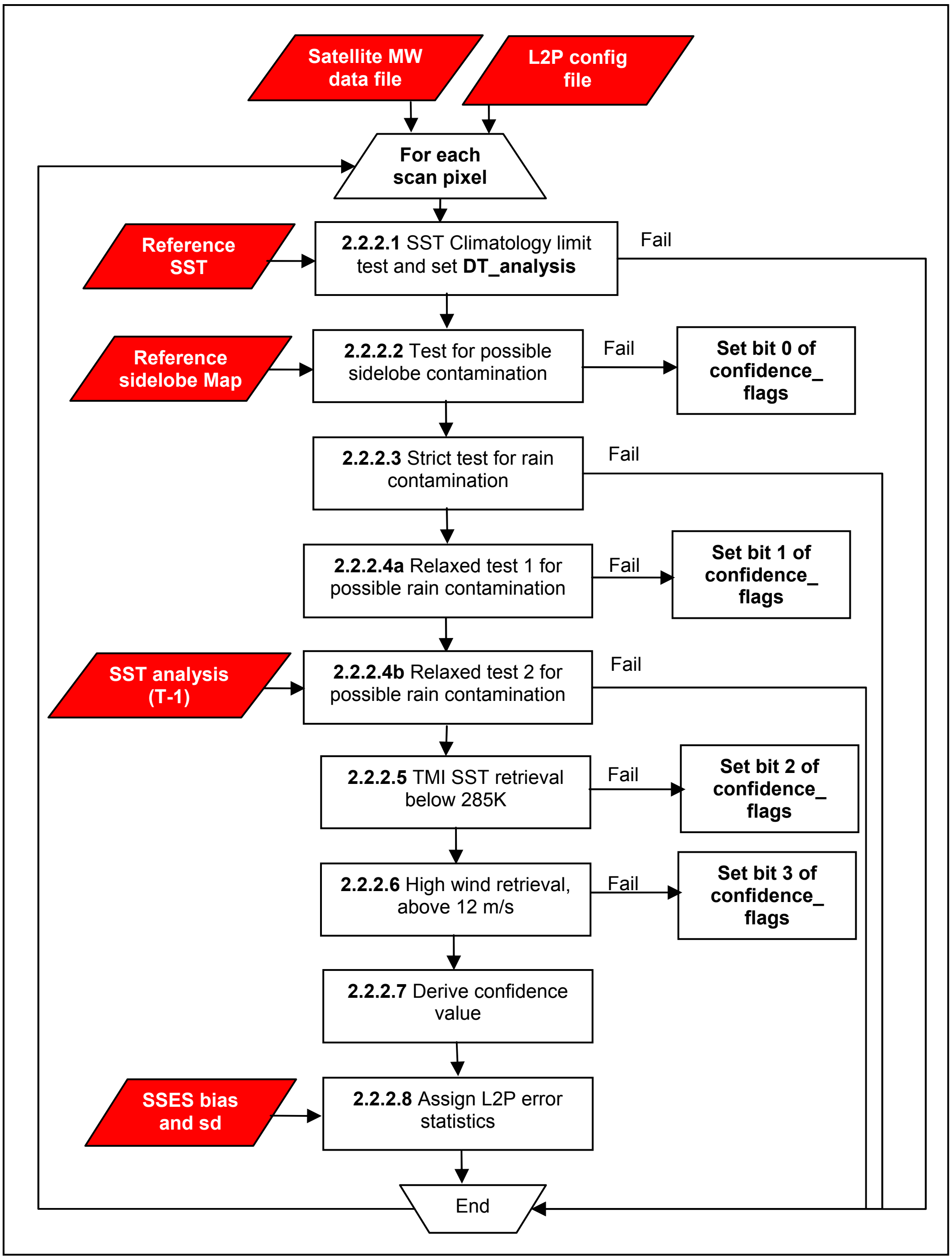


Figure 2.2.2.1 Functional breakdown diagram of data processing steps to produce L2P data records for microwave satellite SST data files.

WP-ID2.2.2.1 SST Climatology limit check

This test is identical to 2.1.1 described above.

WP-ID2.2.2.2 Test for possible sidelobe contamination

Microwave SST observations within about 50 to 100 km from land and sea ice are affected by the warm emission of land and ice entering the antenna side-lobes. The effect varies by region and season (e.g., in monsoon climates where more contamination is expected in wet ground conditions) and is difficult to quantify without having access to engineering data provided by the satellite sensor. It cannot be quantified using buoy match-ups in the coastal zone. Instead, global reference maps identifying regions with persistent sidelobe contamination will be generated by Remote Sensing Systems to provide an estimate the retrieval error in coastal zones and near sea ice margins. The geolocation of each pixel will be used to look up possible contamination from the reference map.

Since the effect is different for the 10.7 GHz and 6.9 GHz retrievals, it must be calculated separately for TMI and AMSR-E radiometers. We will approach this problem in two ways. Initially, TMI will be collocated with VIRS infrared SSTs while AMSR-E will be collocated with MODIS SST retrievals. The bias and variability in the STD will be calculated as the STD as a function of geographic location minus the minimum STD as a function of geographic location. This may be done using only retrievals within 200 km of land or ice or it may be done globally. There are several issues with this methodology that should be considered. There will be some overlap with dependencies on SST and wind speed. If the check is only used close to land, it will ignore possible side-lobe effects near frontal features, but if it is used globally, increased variability from the IR sensors due to water vapour or stray light will be included. It is expected that further work will be required to establish the optimum configuration of sidelobe contamination maps. The GDS specifies the following rule:

- **Rule 2.2.2.2a:** If a pixel is located within an area of known sidelobe contamination as defined by sensor specific sidelobe contamination maps, bit 0 of the L2P confidence data variable **`confidence_flag`** should be set as described in A1.2.3.19.
- **Rule 2.2.2.2b:** If a sidelobe contamination map is not available, any pixel within **`sidelobe_distance_threshold`** (given in km from the nearest land) as specified in the L2P configuration file (see Appendix A9) should be flagged as contaminated.

WP-ID2.2.2.3 Strict test for possible rain contamination

SST cannot be retrieved in raining conditions and data flagged as rain contaminated must be rejected from the L2P data file. The GDS specifies the following rule:

- **Rule 2.2.2.3a:** If a measurement is identified as rain contaminated in the native data input file, this measurement should not be included in a L2P data file.
- **Rule 2.2.2.3b:** If a measurement is rejected due to Rule 2.2.2.3a this should be recorded in bit 3 of the L2P **`rejection_flag`** variable as described in A1.2.3.18 so that a percentage rejection rate can be calculated for the input data file.

WP-ID2.2.2.4 Relaxed test for possible rain contamination

A recognized problem in the MW SST retrievals is unidentified rain contamination in the SST, often nearby raining pixels. Low rain rates or small areas of rain within the large MW footprint are difficult to identify and in many cases, the quality of SST retrievals surrounding a rain flagged pixel are degraded by unresolved rain. For this reason, pixels within a defined radius **`relaxed_MW_rain_distance`** (specified in km and held within the L2P configuration file as defined in Appendix A9) of a raining pixel will be flagged as potentially rain contaminated. In addition, a comparison of the previous analysed SST values will be used to reject or accept the measurement for inclusion within a L2P file if the difference is greater than a predefined threshold specified in K and held in the L2P configuration file variable **`relaxed_rain_threshold`** (see Appendix A9). The GDS specifies the following rules:

- **Rule 2.2.2.4a:** If a pixel lies within a radius of **`relaxed_MW_rain_distance`** specified in km of a clearly flagged rain contaminated pixel, bit 1 of the L2P confidence data variable **`confidence_flag`** should be set as described in A1.2.3.19.

- **Rule 2.2.2.4b:** If the difference between the most recent GHRSST-PP analysis SST and the potentially contaminated pixel is greater than **`relaxed_rain_threshold`** (K), the pixel should not be included in a L2P data file.
- **Rule 2.2.2.4c:** If a measurement is rejected due to Rule 2.2.2.4b this should be recorded in bit 3 of the L2P **`rejection_flag`** variable as described in A1.2.3.18 so that a percentage rejection rate can be calculated for the input data file.

WP-ID2.2.2.5 TMI SST retrieval below 285 K

TMI retrieves SST primarily from the 10.7 GHz channel, which has decreased sensitivity (and therefore increased error) at SSTs below 285K. Although the bias at lower temperatures is negligible, the standard deviation increases (as reflected in the SSES). TMI data having retrievals below 285K must be flagged according to the following GDS rule:

- **Rule 2.2.2.5:** If a TRMM TMI SST measurement value is less than 285K, bit 2 of the L2P confidence data variable **`confidence_flag`** should be set as described in A1.2.3.19.

WP-ID2.2.2.6 High Wind retrieval, above 12 ms^{-1}

Uncertainties in the emissivity model at high winds and increased errors in the directional model, result in larger errors above 12 ms^{-1}. These errors should not vary temporally or spatially since they are due to uncertainty in the estimate of emissivity. Although this increased error will be reflected in the SSES, some data users may chose to simply exclude these retrievals as suspect. The GDS specifies the following rule:

- **Rule 2.2.2.6a:** If a microwave wind speed measurement contemporaneous with a microwave SST measurement is greater than 12 ms^{-1}, bit 3 of the L2P confidence data variable **`confidence_flag`** should be set as described in A1.2.3.19.
- **Rule 2.2.2.6b:** If a contemporaneous microwave wind speed measurement is unavailable, the nearest NWP analysis in space and time to the SST measurement should be used to evaluate Rule 2.2.2.6a.

WP-ID2.2.2.7 Derive a Microwave Proximity Confidence Value (MPCF)

MPCV data are analogous to Infrared Proximity Confidence Values (IPCV) and provide a method to better understand the character of microwave SST and their relationship to the SST1m. Microwave SST measurements are known to be of a reduced quality when they are in close proximity to clearly flagged rainfall, due to side lobe contamination in the proximity of land and sea ice. Also, at wind speeds < 2 ms^{-1}, a cool skin effect may also exist whereas at high wind speeds the emissivity model used to derive SST is poorly defined. These factors may be used to derive a **`proximity_confidence`** value for microwave SST retrievals that can then be used to derive a MPCV for use within the GHRSST-PP MDB and ultimately the assignment of SSES for MW SST data. The initial MPCV scheme proposed for the GDS is based on the following criteria:

1. Physical distance of an SST measurement from the nearest rain flag.
2. Physical distance of an SST measurement from nearest land.
3. Physical distance of an SST measurement from nearest sea ice (held in bit 4 of the L2P **`confidence_flag`**).
4. Conditions when the surface wind speed is > 12 ms^{-1} and is likely to modify the emissivity conditions of the sea surface in the 6-11GHz frequency range.
5. Low wind speed conditions < 2ms^{-1} when a cool skin/warm layer effect may be present.
6. For the special case of TMI SST, conditions when the SST is < 285K.

The proposed scheme is an initial configuration for MPCV and further research is required within the GHRSST-PP to establish an optimal scheme, which should be undertaken at RDAC once sufficient data are available within the GHRSST-PP MDB. The MPCV scale is shown in Table 2.2.2.7.1.

Table 2.2.2.7.1 Microwave proximity confidence value (MPCV) scale definitions.

MPCV Value	Description
10	**Unprocessed:** Data that have not been classified (measurement indicates sea ice)
11	**Questionable:** Data that are may be contaminated by land, rain, sea ice, RF interference, degraded due to uncertainty in seawater emissivity at higher wind speeds, and/or are below a low SST threshold.
12	**Acceptable:** Data that are far from land, rain, sea ice, RF interference and are within a favourable wind speed regime.
13	**Diurnal:** Data are far from any rain or land flags but are within a low wind speed regime where a cool skin or warm layer may influence the observed SST.

If any of the test bitflags are set for the MW retrieval, the data will be flagged as questionable. In addition two surface wind speed threshold values are used to delineate the wind speed conditions for which the best SST retrievals and relationship to the in situ observations is expected using the threshold parameters **`MPCV_wind1`** and **`MPCV_wind2`.** The values for **`MPCV_windn`** shown in Figure 2.2.2.7.1 are expected to vary for a given sensor and the exact value of these parameters is a subject of on-going research and development within the GHRSST-PP. Finally, the physical distance in km specified in the L2P configuration file variable **`MPCV_proximity`** is used to identify pixels that are far away from contaminating effects of sea ice (flagged in WP-ID2.1.3) as described in Appendix A9.

All threshold values will be stored in a L2P configuration file specific to each RDAC or GDAC and maintained at the GHRSST-PP international project office that can be modified based on an assessment of sensor specific values (see Appendix A9). This will only be possible once the GHRSST-PP MDB has sufficient data available for analysis. The filename of this configuration file should be written into each L2P file using the optional global netCDF attribute "References" as described in Appendix A1.2.1. It is expected that the methodology used to derive MPCV will be modified based on experience in the GDSv2.

Wind speed value (ms^{-1})	Any bit of confidence_flag has been set		
`u < MPCV_wind1`	11	13 (Diurnal)	
`MPCV_wind1 < u < MPCV_wind2`	11 (Questionable)	12 (Acceptable)	
u > MPCV_wind2	11	11	
		> `MPCV_proximity` (far)	**Distance from nearest problem area (km)**

Figure 2.2.2.7.1 Schematic diagram showing the initial scheme (GDS-v1.0) proposed to derive MPCV values. Note that the TMI is subject to an additional lookup based on the actual SST value which is not shown.

WP-ID2.2.2.8 Assign L2P Sensor Specific Error Statistics

The following GDS rules are used to specify the MPCV value in L2P data records:

- **Rule 2.2.2.8a:** If pixel bias error and sd error statistics are provided by a data provider, these values may be assigned to the L2P **`sses_bias_error`** and **`sses_standard_deviation_error`** variables respectively as described in A1.2.3.3 and A1.2.3.4. In this case, bit 6 of the L2P **`confidence_flag`** variable should be set as described in A1.2.3.19 because these error estimates are independent of GHRSST-PP methods.

- **Rule 2.2.2.8b**: If bit 6 of the L2P **`confidence_flag`** variable has been set, it is strongly advised that the associated **`proximity_confidence`** value is also updated as described in A1.2.3.20 by the data provider. The values should be fully compliant with the GDS MPCV scale. In addition, bit 7 of the L2P **`confidence_flag`** should be set as described in A1.2.3.19 to avoid inconsistency.

If native **`sses_bias_error`** and **`sses_standard_deviation_error`** variables values are not available or an RDAC chooses to ignore the error statistics provided by a data provider, these must be assigned based on the L2P **`proximity_confidence`** variable using sensor specific error statistics (SSES) as described in WP-ID3. The GDS specifies the following rule:

- **Rule 2.2.2.8c:** If pixel **`sses_bias_error`** and **`sses_standard_deviation_error`** error statistics are not available for SST measurements, these values should be assigned the most recent SSES bias and sd values for the each data stream and L2P confidence data record **`proximity_confidence`** value. In addition, bit 6 and bit 7 of the L2P **`confidence_flag`** should be set to zero as described in A1.2.3.19.

It is expected that the MPCV and SSES schemes will develop considerably throughout the GHRSST-PP.

WP-ID2.3 Evaluate L2P data

Each RDAC and GDAC centre will evaluate whether or not a L2P data set should be admitted to the GDS processor based on objective criteria Including completeness, timeliness and quality. Statistics for rejection assessment may be computed using the L2P **`rejection_flag`** variable as described in A1.2.3.18. Internal rejection variables may be required for other data sources (even auxiliary data sources) to maintain records of file rejection rates.

WP-ID2.3.1 Reject L2P data file

If a L2P data file has been rejected by an RDAC/GDAC, an error message must be raised and logged at the GHRSST-PP ERRLOG system as described in Appendix A7.2. No MMR_FR metadata are prepared for rejected L2P data streams.

WP-ID2.3.2 Accept L2P data file

If a L2P data file is accepted for further use in the GDS, then a L2P data file should be generated as described in WP-ID2.4. A corresponding MMR_FR metadata record should be generated as specified in WP-ID2.5.

WP-ID2.4 Format L2P data file

Each L2P data set should be prepared, formatted and archived as a netCDF file following the data format described in Appendix A1.2 and A1.4.

WP-ID2.5 Generate and register L2P MMR_FR metadata record

For each L2P data a corresponding MMR_FR must be prepared by the responsible RDAC/GDAC processor and registered at the MMR as soon as the L2P data become available to the user community.

WP-ID2.5.1 Format L2P MMR_FR metadata record

A MMR_FR metadata record should be prepared for each L2P data file according to the specification provided in Appendix A6.2.

WP-ID2.5.2 Register L2P MMR_FR metadata record with MMR system

MMR_FR metadata records should be delivered to the MMR system as soon as possible after the L2P data set has been evaluated following the procedure described in Appendix A6.4. RDAC should aim to register MMR_FR records within 60 minutes[11] of L2P data production in order that users are informed in a timely manner.

[11] It is recognised that if the MMR_FR record is submitted by e-mail delays in registration are beyond the control of the RDAC. At the 4th GHRSST-PP Science Team workshop, it was proposed to enable MMR data records to be delivered to the MMR via ftp push from the RDAC to circumvent this problem.

WP-ID3 The GHRSST-PP Matchup Data base (MDB) and derivation of Sensor Specific Error Statistics (SSES)

WP-ID3	Derivation of Sensor Specific Error Statistics (SSES)
Work Package number :	**WP-ID3**
Leader:	**GHRSST-PP Science Team and GHRSST-PP Project Office**
Aim:	**To specify the procedures and methods that will be used to derive Matchup Database (MDB) records and Sensor Specific error statistics (SSES) for each input satellite SST data stream.**

1	**Objectives:** 1. To specify the space and time constraints to match in situ SST measurements to satellite SST measurements. 2. To specify the format of GHRSST-PP validation records that collectively form the GHRSST-PP match up data base (MDB). 3. To specify the procedure for generating a MDB data record 4. To specify the procedure for delivering MDB records to the GHRSST-PP MDB. 5. To specify the analysis methods that will be used to derive Sensor Specific Error Statistics (SSES) based on MDB data records.
2	**Description:** This WP describes the GHRSST-PP match up database system (MDB) and the MDB data records that should be produced at each RDAC. MDB records contain in situ SST (and other ocean-atmosphere data) that are matched to L2P satellite SST measurements according to specified time and space constraints. The GDS requires that error estimates are assigned to every satellite SST measurement during the derivation of L2P data products. As an in situ observation is not available for every SST measurement, Sensor Specific Error Statistics (SSES) are derived by a statistical analysis of MDB data records for a given sensor type and time period. SSES consist of a statistical mean bias and standard deviation (sd.) error estimate that is correlated to the mean L2P **proximity_confidence** value for a given sensor. These values are then assigned to all L2P measurements that have the same **proximity_confidence** value. In this way, optimal use of the limited amount of in situ data can be made. A precedent for this approach has been set by the EUMETSAT O&SI SAF and the US Navy NAVOCEANO operational SST data production systems as discussed at the third GHRSST-PP workshop. In addition, The MDB system also provides a resource that can be used to validate GHRSST-PP data products. In the case of microwave SST, the derivation of SSES also requires that error components are assigned to each pixel based on static look-up-tables of mean error due to sidelobe contamination, proximity to rain and wind speed. These tables will be provided by Remote Sensing Systems (RSS) for the TMI and AMSR-E (and in the future, Windsat) microwave instruments. The main source of in situ observations used in this system will be those available from the GTS supplied through international in situ data centres (e.g., MED, CORIOLIS, FNMOC) at which QC procedures are used to filter out poor quality data. These data centres only provide in situ observations in a delayed mode (typically 24-72 hours, although delays may be longer than this) and therefore RDAC and GDAC can only provide MDB records in a delayed mode. RDAC are expected to arrange access to regional in situ data sources that are not available via the GTS system but can be included in the MDB system. The GHRSST-PP reanalysis project will be able to make use of other in situ observations that are unavailable in a NRT mode.

3	**Inputs:** 1. L2P data files 2. In situ ocean-atmosphere measurements from buoys, drifters and ships 3. SSES configuration file 4. Static Look-up-Table (LUT) of microwave SST sidelobe contamination error 5. Static LUT of microwave SST rain contamination errors 6. Static LUT of microwave SST wind speed errors
4	**Outputs:** 1. MDB records delivered to the GHRSST-PP MDB 2. SSES for each sensor derived on a on a regular basis
5	**Acceptance tests:** 1. MDB records are formatted correctly and are delivered to the GHRSST-PP MDB system 2. SSES analysis is performed on a weekly basis and SSES are published to all RDAC and via the GHRSST-PP UIS
6	**Metric for performance assessment:** 1. MDB records are produced in a timely manner using a target of within 24 hours from when the in situ data are made available by a data centre. 2. MDB records are provided and ingested at the MDB in near real time. MDB rejection rate < 2%

This work package is dedicated to the generation of satellite SST matched to near contemporaneous in situ observations and stored in the GHRSST-PP Matchup Database (MDB). The MDB will be physically realised as a complete relational database system that can be analysed using structured query language (SQL) routines. These data provide a collective resource that is shared and populated by all RDAC and GDAC within the GHRSST-PP to generate an independent assessment of the absolute error of all satellite SST data streams. The MDB forms the core data resource for both the derivation of Sensor Specific error statistics (SSES) that are assigned to all L2P data in WP-ID2 and for the on-going validation of satellite data streams described in WP-ID5. Both of these rely on the regular population of the GHRSST-PP MDB using near real time satellite and in situ observations.

Figure 3.1.1 presents a functional breakdown diagram of the processing tasks within WP-ID3 that is split into two interrelated components:

1. The generation of individual GHRSST-PP MDB records and their storage in the GHRSST-PP MDB database system
2. The analysis of MDB data records to derive Sensor Specific Error Statistics.

In summary, GHRSST-PP L2P data files for each individual satellite sensor are matched in space and in time with in situ observations (WP-ID3.1.1) if available. Time and space constraints that define near contemporaneous match-up criteria are stored in an MDB configuration file which will be maintained by the GHRSST-PP International Project Office. If in situ data can be matched with satellite data then a MDB data record is prepared and formatted according to the specification set out in Appendix A4.1 (WP-ID3.1.2). Note that in situ observations may be matched to SST measurements derived from different satellite sensors and are thus used more than once in the system providing consistency across GHRSST-PP data products. However, only 1 instance of each in situ observation (the nearest in space and time to the satellite data) is used to define a matchup record.

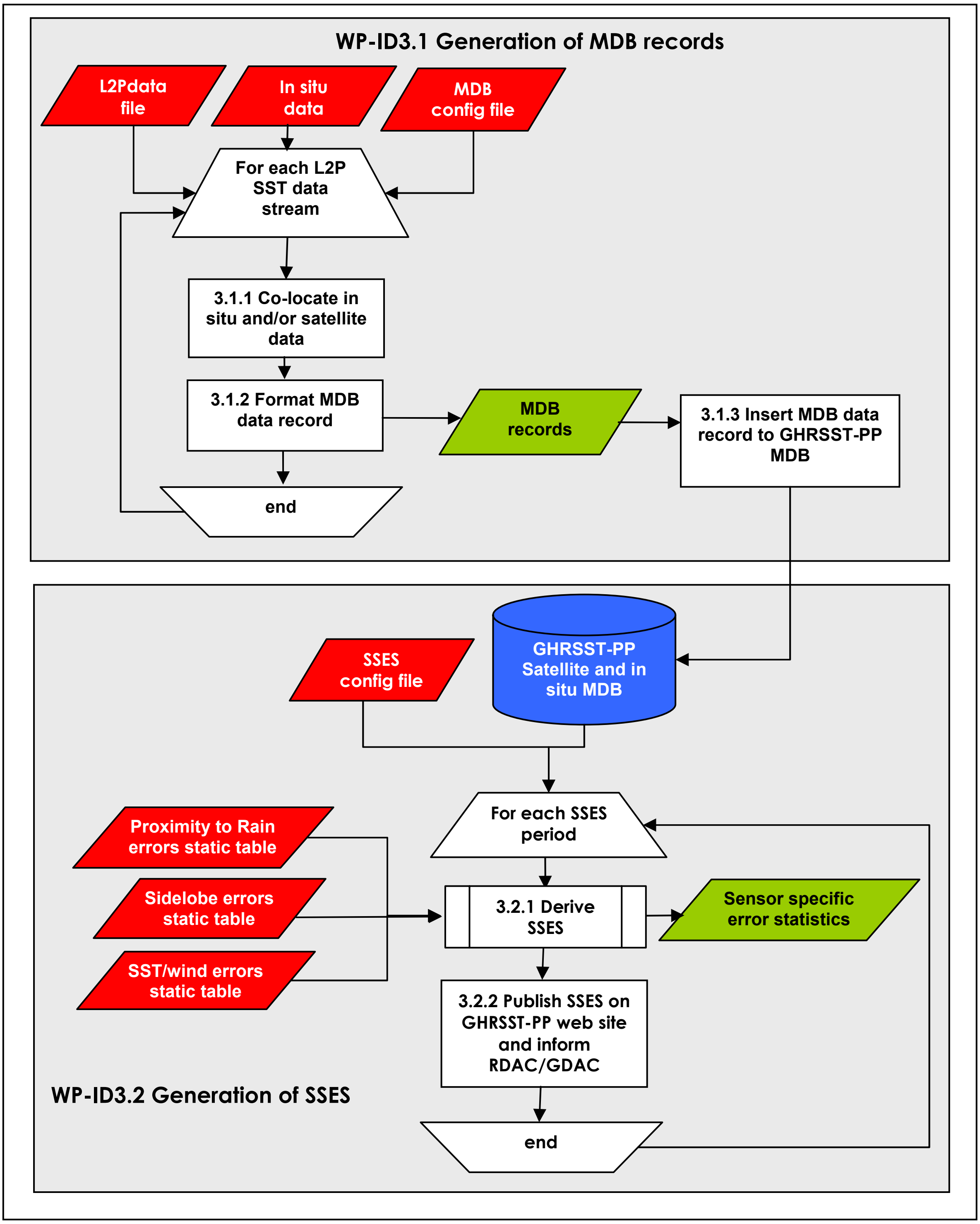


Figure 3.1.1 Functional breakdown diagram of WP-ID3.

At regular intervals, known as SSES periods, an analysis procedure is used to derive time-average bias and sd. value for each satellite sensor and for each L2P `proximity_confidence` value (WP-ID2.2). SSES provide independent error quantification for each L2P data stream. SSES are published on the GHRSST-PP web site as part of the GHRSST-PP User Information Service (UIS) and all RDAC/GDAC are informed of the latest issue of SSES for use in the generation of GHRSST-PP L2P data products. A SSES configuration file is used to control the derivation of SSES statistics for each L2P data stream which will be maintained by the GHRSST-PP International Project Office and available at the MDB itself.

WP-ID3.1 Generation of Match Up Database (MDB) records

The GHRSST-PP MDB system is a relational database containing all GHRSST-PP MDB data records. The GHRSST-PP MDB will be implemented and maintained at the JPL PO.DAAC using the MySQL /database system. Access to the MDB (either a web based user interface or remote command shell) will be unrestricted to all participants of the GHRSST-PP but will be subject to the requirements and stipulations of data providers (in some situations, match up data records are sensitive data *e.g.,* shortly after the launch and commissioning of a satellite sensor). General access to the MDB will require that users are registered and agree to any data access and usage agreements that are relevant to MDB records. Structured Query Language (SQL) statements will provide simple yet extremely powerful analysis routines using all or a just a sub-set of MDB records maintained within the database system itself.

MDB data records should be produced by RDAC and ingested at the MDB in near real time (1-7 days[12]) in order for SSES to be generated in a timely manner. Timeliness of MDB data is related to the availability of in situ data. Time penalties are incurred because data centres careful quality control and validate in situ measurements. Collaboration with operational data centres that are well practiced in the QC of in situ observations must be established and maintained by RDAC and GDAC especially in the case of data that are made available on the GTS. The GHRSST-PP expects to work closely with the CORIOLIS data centre (France), the Japanese Meteorological Agency (JMA), and the MEDS data centre (Canada) amongst others as described in Appendix A3.4.5 and Appendix A3.4.6. However, the MDB is open to all in situ observations that are deemed of sufficient quality by the RDAC and GDAC teams.

WP-ID3.1.1 Derivation of MDB data records

The data inputs to GHRSST-PP MDB records are L2P data streams and in situ observations. For an MDB data record to be accurate and legitimate the comparison must be between like measurements although in the case of SSTskin observations, in situ data are extremely scarce. Instead, emphasis is placed on the use of in situ SST1m data derived from drifting/moored buoy and ship observations. These data provide a reference data source to which all satellite data are referred to when deriving bias estimates. While the use of SST1m data in this way is not as appropriate as SSTskin observations in some cases, the small number of in situ SSTskin observations limits the relevance of such data within the GDS. However, the use of SST1m data is extremely appropriate to the climate SST data record which is also referenced to SST1m. It is encouraged to consider all available in situ data down to 100 meters in the MDB records generation.

In situ SST data are provided from a number of different sources and will be of variable quality. Some data providers provide high quality QC'd in situ data streams in near real time whereas in situ data drawn down directly from the GTS must be QC'd locally before use. Other data streams are only available to national institutes where they must be pre-processed and matched to satellite data on a case by case basis. The great diversity of in situ data sources and quality make it difficult to specify exact rules for the QC of these data and each RDAC and GDAC centre must be responsible for implementing basic in situ QC procedures using local ocean/atmosphere knowledge and experience.

Since satellite-to-in situ data are typically spatially and temporally inhomogeneous (leading to the term 'near contemporaneous'), this implies that limits of coincidence must be applied to constrain the contributions to the error budget of the SSES/validation procedure from these sources to acceptable levels (say <0.1K). In the case of in situ-to-satellite data match ups, the following GDS rules are specified:

[12] Note that the MDB will be used extensively by the GHRSST-PP reanalysis project so that the time constraints specified here should be interpreted in the broadest possible sense; if a match-up data record has been established, it should be delivered to the MDB system regardless of the timeliness of the data.

- **Rule 3.1.2a:** Ideally, the spatial separation of the validating measurement and the centre point of a satellite-based measurement (centre of pixel) should be 2-3 times the standard deviation of a well mapped satellite image. In the case of high resolution (~1km at nadir) data streams considering a maximum geolocation error, this is approximately 3 km. However, it is recognised that many operational centres already have systems that use less stringent criteria (e.g., NAVOCEANO use 25 km and within 6 hours) to define a match up with little impact on the overall statistical results following analysis. By using more relaxed criteria to define match-up coincidence, the number of match-up data records is significantly increased and it is always possible to filter these data and generate a sub-set of these records to more stringent criteria if required.

  In the case of low resolution gridded microwave SST measurements (grid size of 25 km) clearly the spatial stringency for match up criteria do not apply. The mean of a transect of several measurements obtained within the grid cell will provide a better representation of the in situ SST and are preferred to single point observations from moored or drifting buoys. The GHRSST-PP match-up criteria[13], which should be applied to high resolution and gridded coarse resolution satellite data streams, are maintained in the MDB configuration file using the variables **`MDB_time_constraint`** and **`MDB_space_constraint`** of the MDB configuration file specified in A9.7. In situ data that cannot be matched to better than these criteria should be ignored. Once sufficient data have been ingested by the MDB, analysis of all records will define the acceptable upper limits for these situations. As the GHRSST-PP MDB system will be a relational database system in which it is a relatively trivial task to select data with a more constrained matchup criteria specification in both time and space, the GDS specifies **`MDB_time_constraint=6 hours`** and **`MDB_space_constraint=25km.`**
- **Rule 3.1.2b:** Only one instance of each in situ observation for each individual satellite sensor should occur within the MDB i.e. the satellite data having the smallest deviation in space and time from the in situ observation should be used within the MDB and all other duplicates should be removed.
- **Rule 3.1.2c:** L2P data adjacent to the validation measurement must be extracted to provide information on the spatial variability in the vicinity of the validation point. Appropriate 'missing data' values will be used if any of the elements of the pixel array extend beyond the edges of the instrument swath. A 5 x 5 array of L2P data should be extracted centred on the validation pixel. The 5 x 5 box array is selected as a valid MDB record only if therfe is at least 1 valid SST satellite SST measurement within 25km of the in situ measurement.
- **Rule 3.1.2d:** All L2P confidence data each pixel stored in a MDB record must accompany the extracted data.
- **Rule 3.1.2e**: The timeliness of the validating data is important, but speed of delivery should not be used as an excuse to compromise the quality of the measurements. A minimum delay of 24 hours is to be expected when using data from operational in situ data centres due to the latency incurred by QC procedures. However, all MDB records are admissible irrespective of their timeliness for use by the GHRSST-PP reanalysis project.
- **Rule 3.1.2f:** Additional geophysical fields will be required to provide insight into any uncorrected perturbations to the SSTs which might result in systematic errors. All available in situ parameters should be included in the MDB record using the spare fields that have been made available for this purpose.
- **Rule 3.1.2g:** When matching satellite data to in situ data, Temporal coincidence takes precedence over spatial coincidence.

[13] Based on discussions at the 4th GHRSST-PP Science Team Workshop, Los Angeles, USA, September 2003, **`MDB_time_constraint`** = 25 km radius of the pixel centre point and **`MDB_space_constraint`** should be within ±6 hours of the satellite overpass. These constraints are considered the absolute maximum and all efforts should be made to reduce these to the smallest deviations from zero as possible.

WP-ID3.1.2 Format of MDB record

Each in situ SST matched to a near contemporaneous satellite SST will be formatted as an XML text file following the format described in Appendix A4.

WP-ID3.1.3 Insert MDB record into the GHRSST-PP MDB system.

MDB data records should be sent to the GHRSST-PP MDB either as an e-mail message or via ftp put. Messages/ftp files will be parsed and inserted into the GHRSST-PP MDB if correctly formatted. An error message will be returned to the data provider if the MDB record is incorrectly formatted.

WP-ID3.2 Derivation of sensor specific error statistics (SSES) using MDB records.

Satellite SST data sets are have errors that vary in time and space due to a number of different reasons as shown in Figure 3.2.1. Idealised temporal bias and standard deviation (shown as error bars) estimates are shown in this figure for five separate satellite data streams that should be viewed as the product of many statistical matchups between in situ SST data that are contemporaneous with the satellite data. Some data sets overlap in time with each other allowing sensor inter/cross-calibration whereas others only have a minimum (if any) temporal overlap. The data could have been derived from either microwave or infrared sensors could be polar orbiting, geostationary or low-earth orbit satellite platforms.

The first data set, shown in light blue to the far left of Figure 3.2.1, shows an infrared sensor that has an improving bias error but a degrading standard deviation error. This could be a real effect caused by the sensor/satellite system entering normal operations following a commissioning period during which system tests were influencing the error statistics. A more likely explanation for the improvement in mean bias at the cost of an increase in standard deviation is more likely to be a function of less stringent spatial and temporal criteria that are used to match the satellite to in situ data. The data set is then abruptly degraded following the injection of stratospheric aerosols caused by a large volcanic eruption. Both mean bias and standard deviation increase ands then decay initially quite quickly as the aerosol are dispersed, but then more slowly as the smaller aerosol have much longer residence times.

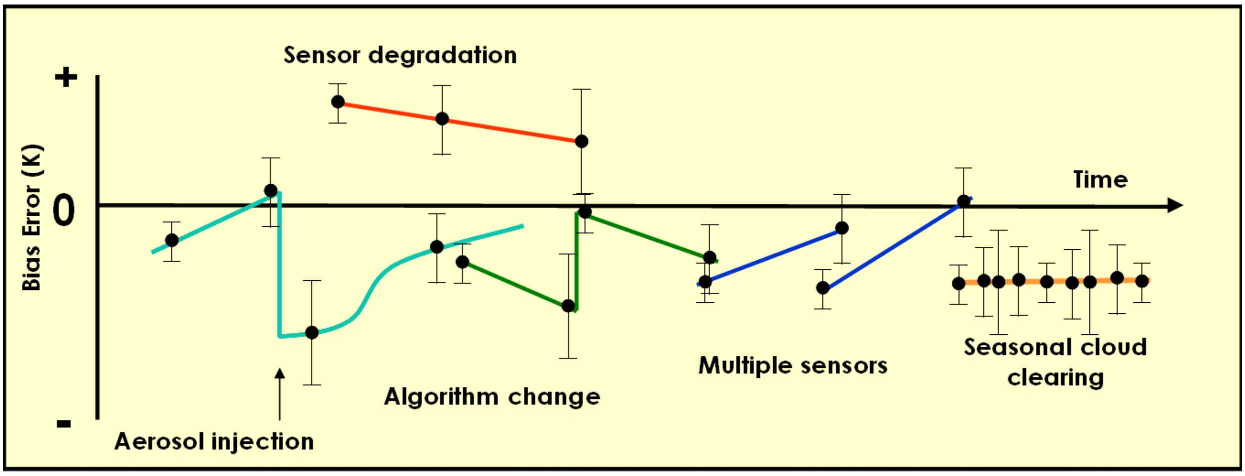


Figure 3.2.1 Schematic diagram showing how the satellite SST data record has a time varying bias and standard deviation due to a variety of different causes.

The second example data set shown in red on Figure 3.2.1 shows the effect of sensor degradation over time. In this particular example, while the sensor is degrading (oxidation of sensor components for example) there is an increase in standard deviation but the bias error is reduced over time as the atmospheric aerosol injected earlier on in the record continually decays. Of course, degradation could equally cause an increase in bias and standard deviation due over time.

Example three in Figure 3.2.1 (shown in green) shows a typical pattern of error associated with an algorithm change. Bias and standard deviation error initially increase and are abruptly reduced following a change in algorithm, only to increase again. This pattern may be associated with an on-board calibration error (e.g., a slow drift of calibration reference targets) or an SST algorithm that is unable to account for spatial and/or temporal variations in atmospheric structure. An alternative explanation could be that more in situ data have been made available for validation or that the in situ data are of a higher quality (e.g., a larger than average number of research cruises). The bias error is better characterised using the newly available in situ observations but degrades as the in situ database returns to levels similar to the start of the mission. In this scenario, the satellite sensor is performing extremely well and has a minimal bias and standard deviation but the lack of sufficient high-quality in situ data is the limiting the actual quantification of the errors.

The fourth example shown as blue lines in Figure 3.2.1 shows two well calibrated sensors that have steadily improving bias and standard deviation errors due to (for example) a combination of improved SST and sensor algorithms and better in situ validation campaigns. Considerable overlap between a follow on mission has been planned as part of the measurement strategy providing continuity across the two independent data sets. This is essential if satellite data are to be used together to provide high quality climate data records.

The final example shown in Figure 3.1.2 (orange line) shows a particularly interesting case in which there is a static time invariant bias error (for example a sensor that is providing an estimate of SSTskin but is validated using near contemporaneous SSTdepth measurements) but a seasonal standard deviation error associated with seasonal variations in cloud cover. In this example, the influence of clouds on an infrared data stream could cause an increase of sub-pixel cloud contamination or a reduction in the number of validation data available for analysis or a combination of both.

Figure 3.2.1 is meant to provide an illustration of the complexity involved in establishing error estimates for satellite SST data streams rather than an exhaustive litany of possible causes of error. It is clear from the discussion that the choice of reference data from which bias and standard deviation errors are derived is critical. The reference data set should be accurate, temporally stable, provide global coverage and be available for all sensors at all overpass times. This is a demanding set of requirements!

There are considerable advantages to using satellite data as a reference SST in terms of coverage, consistency, calibration and accuracy and timeliness. Unfortunately, the main risk is that if the reference satellite sensor were to fail, a follow on may not be available immediately and it may not be possible to inter-calibrate a surrogate to the same standard. Additional complications arise due to view geometry differences, spectral differences and the fact that true collocation between different sensor data streams will be limited – especially in the mid and low latitude regions. A satellite SST could be used together with the in situ array if it can be demonstrated that the satellite SST is stable and accurate enough in all parts of the ocean.

Instead, the GHRSST-PP Science Team have chosen to adopt SST1m measured by drifting/moored buoys and ships as a reference SST for all satellite SST data used within the GDS. The main reasons for choosing this data stream above all others are:

1. The availability of in situ infrastructure providing measurements of SSTskin or SSTsub-skin is extremely limited and far from operationally viable (initial trials of autonomous in situ radiometer systems have only recently begun).
2. There is an extensive array of SST1m operational observations covering most ocean areas (the most notable exception is the Southern Ocean). There is a considerable

amount of operational experience using these data which are reported in real time via the GTS system. Several data centres provide GTS data that have been quality controlled incurring short delays of 24-48 hours.
3. There is a demonstrated history of using in situ SST1m for the validation of satellite SST data sets.
4. Most contemporary ocean modelling systems require a measure of the SST1m rather than the SSTskin or SSTsub-skin.
5. Climate SST data sets are based on the SST1m. Furthermore, satellite SST data must be referred to SST1m if they are to enter the climate record.

The calibration of in situ SST1m data remains an issue as it is difficult in many cases to perform both a pre- and post-deployment calibration as the buoy is typically lost. Moored buoys although limited in number, have better calibration accuracy as the sensors are more easily maintained. Ship observations are generally of poorer quality due to warming by pump gear, warming by engine room temperature differentials and un-insulated pipe work between the temperature sensor and water intake aperture within the ship. In addition, as ships are loaded differently on each voyage, the depth at which the SST is measured is variable (See Emery et al, 1997). In general the accuracy of moored buoys is better than 0.1K, drifting buoys 0.1-0.2 K and ships 0.1-1.5K.

In order to establish a bias error and standard deviation for each satellite SST data stream, a sufficient density of in situ SST1m data must be available for each overpass to provide meaningful statistics. Ideally, for each spatial area of 12° x 12°, covering the global ocean, at least 2 buoy in situ measurements (or 6 ship observations[14]) are required for at least 75% of weeks in a given season to provide an adequate satellite SST bias correction prior to their use in optimal interpolation systems (Reynolds and Zhang, 2003). The GHRSST-PP will provide SST data products every day rather than the weekly timescale discussed here and intend to provide error statistics for each satellite measurement. Consequently there are insufficient in situ data available to provide a complete daily bias correction. In order to circumvent this problem, the GHRSST-PP Science Team has agreed to use a time average statistical bias and standard deviation computed for each specific sensor.

Rather than simply compute an overall mean bias for each data stream and apply this value to all data originating from a given sensor, the GHRSST-PP intends to compute error statistics that are refined based on a number of objective criteria that are known to increase bias error. For example in the case of infrared SST data, it is known that in the immediate vicinity of cloudy areas sub-pixel clouds are often present but cannot be detected. This and other effects, such as atmospheric aerosol for example, can significantly increase the bias error. In the case of microwave SST, sidelobe contamination and limitations of the surface emissivity model at high wind speed are known to increase errors. These effects are manifest in the **`proximity_confidence`** values assigned to each satellite SST measurement in WP-ID2. As errors are specific to each particular sensor, these must be computed separately for each sensor leading to the concept of Sensor Specific Error Statistics (SSES) which is the method used by the GHRSST-PP to assign bias and standard deviation error statistics to each satellite measurement based on their **`proximity_confidence`** values.

WP-ID3.2.1 Computation of SSES using MDB records

At regular SSES periods (weekly, monthly or quarterly for example) SSES will be derived for each **proximity_confidence** value and each of satellite SST data stream indicated in Appendix A3,

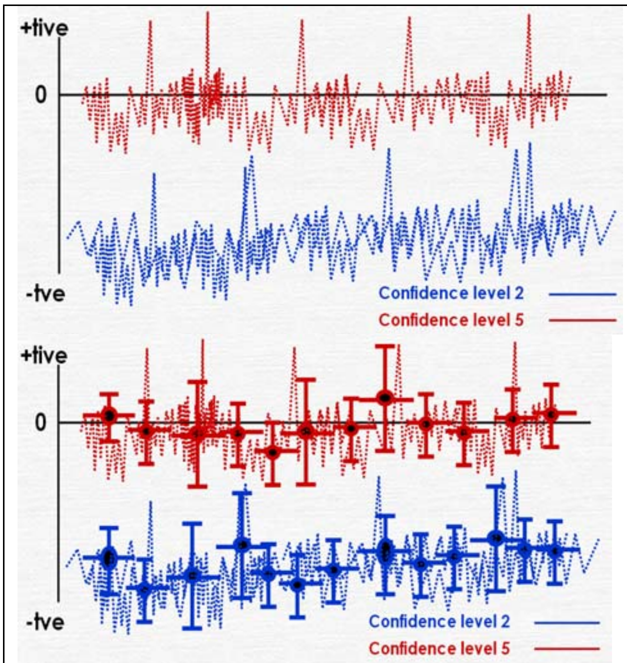
[14] Following Reynolds and Zhang (2003), assuming that the error used in an optimal interpolation scheme for each buoy is 0.5K and that the error for a ship observation is 1.3K. Further, assume that the ship error is random, then the number, *n*, of ship observations to reduce ship error to 0.5K is $1.3/\sqrt{n} = 0.5$. Thus, 6 ship observations roughly equal 1 buoy observation.

Table A3.2. SSES are input into the GDS in WP-ID2 (generation of L2P data products) where a SSES value is assigned to each input SST measurement based on the **proximity_confidence** value that has been assigned.

SSES are derived using an analysis of MDB records in the following way:

1. All MDB records for a given sensor and SSES period are extracted from the GHRSST-PP MDB system.
2. Data are stratified according to their **proximity_confidence** value.
3. The data are then analysed by computing the mean bias and sd. (Satellite minus in situ) for each **proximity_confidence** value. In the case where insufficient data are available within a given SSES period to calculate a significant mean and sd., data from previous SSES periods may be used.
4. The SSES values computed from a limited number of matchup data held within the MDB can then be assigned to all other measurements having the same **proximity_confidence** value throughout the SSES period.

Through the computation of SSES, the limited numbers of in situ data are used in an optimal manner.

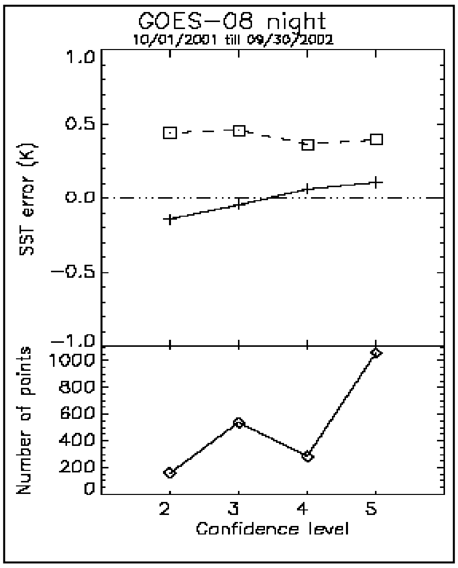


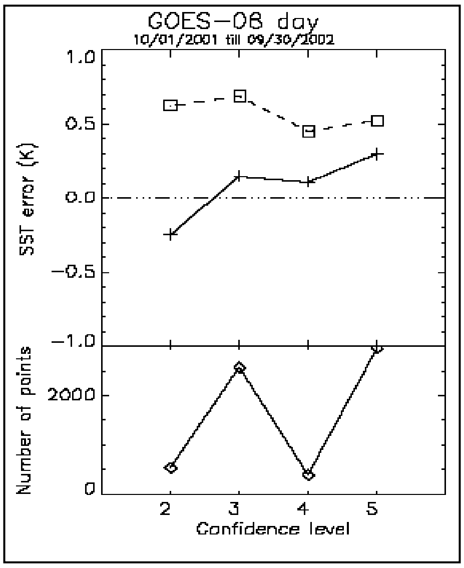
(a)	(b)

Figure 3.2.2 Schematic diagram of SSES for a hypothetical sensor. (a) Time series of SST bias error computed for all available data within the MDB for Proximity confidence values 2 and 5. (b) SSES bias error and standard deviation computed for SSES periods plotted over the MDB data.

Figure 3.2.2 shows schematically the derivation of SSES. In panel (a) a time series of all bias error data extracted from the GHRSST-PP MDB and stratified according to **proximity_confidence** value have been plotted for **proximity_confidence** (PC) values 2 and 5. PC value 5 data (the highest confidence) have a smaller bias error compared to confidence level 2 data. Note that during certain periods, the density of matchup data within the MDB is much higher. Panel (b) shows the SSES mean bias and associated standard deviation calculated for consecutive SSES period.

Figure 3.2.3 provides an example of SSES for GOES-8 SST computed CMS, France over the period 10/2001 to 10/2002. In this example separate SSES have been computed for day and night time conditions. The relationship between **proximity_confidence** value and mean bias is readily seen in this figure which treats day time and night time data differently. Figure 3.2.4 shows an example application of the scheme to a GOES image in the Atlantic Ocean. It is clear from this Figure that the use of confidence flagging provides additional quality control in formation for areas that would otherwise be considered good SST retrievals (e.g., the area off the coast north of the River Plate estuary, Uruguay/Brazil).





(a) (b)

Figure 3.2.3 SSES errors derived for GOES-8 SST processed at the O&SI SAF. Crosses indicate bias error and open squares indicate SD using a 5 point proximity confidence scale. (a) Night-time data (b) Daytime data. Confidence and SSES are derived from a dedicated GOES-8 MDB based at CMS, France over the period 10/2001->10/2002 using a match-up criteria of 5x5 IR pixels. (P. LeBorgne)

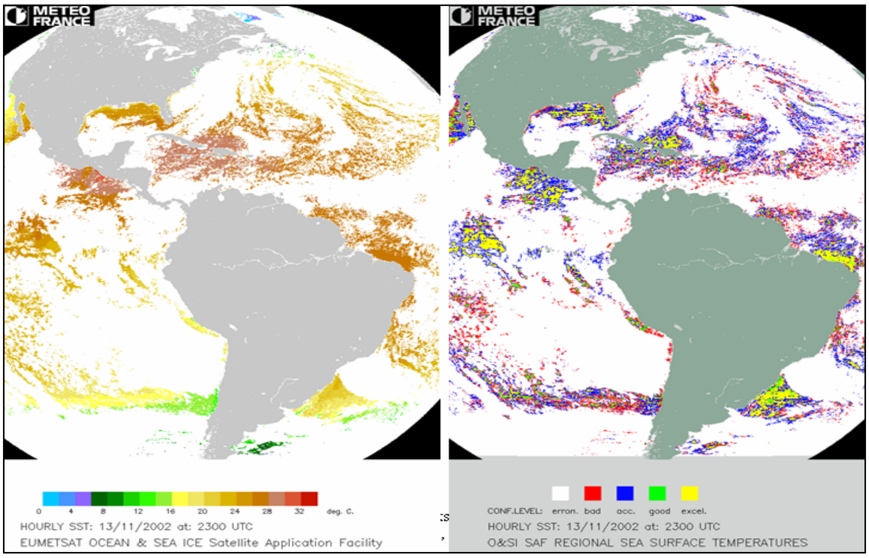


Figure 3.2.4 (Left) GOES MCSST for 13/11/2002 at 23:00 UTC. (Right) confidence value for the image data shown in panel (a). (P. LeBorgne)

WP-ID3.2.2 SSES for microwave SST data

The derivation of SSES for microwave SST data is more complex than for infrared SST data. This is already apparent from the difference between methodology used to derive proximity confidence values for microwave (MWPC) and infrared (IPCV) SST data streams. The retrieval error for microwave SST depends on the actual radiometer noise-equivalent delta temperature (NEDT) for all measurements used in the SST retrieval, the environmental characteristics at the time of measurement (since the surface emissivity in the microwave frequency is primarily a function of SST and wind speed), and other factors such as proximity to land and sea ice. Errors due to diurnal warming are not included in the following discussion but should also be considered as these relate to the reliability of the MDB data themselves. Collocations between noon and 4PM, with wind speeds of less than 4 m/s are almost certainly affected by a significant diurnal warming effect and a consequent decoupling of SST1m data from satellite derived SST.

For microwave SSES bias calculation, a daily mean bias will be calculated from the GHRSST-PP MDB data records. At Remote Sensing Systems (Primary near real-time data provider for TMI and AMSR-E SST data sets) near real-time (NRT) bias corrections are performed using NRT in situ observations that are downloaded daily from the Global Ocean Data Assimilation Experiment (GODAE) Monterey server, which is sponsored by the Office of Naval Research (ONR) and hosted by the Fleet Numerical Meteorology and Oceanography Center (FNMOC). These observations are obtained by FNMOC from the GTS and processed for the GODAE server. Observations from ship engine room intake, fixed buoy, drifting buoy, ship hull sensors, and CMAN stations are included in the dataset. The US-GODAE website (http://www.usgodae.fnmoc.navy.mil) has the complete SURFOBS dataset and a detailed description.

In situ observations are collocated with the closest AMSR-E or TMI SST observations. A collocation is made only if there is a satellite observation within 25 km and 6 hours. Gentemann (accepted JGR) showed that the NDBC buoy array had much larger biases and standard deviations than the open-ocean TAO and PIRATA arrays. The NDBC array is compromised of buoys located in the coastal US waters, and are often located quite near to land. Therefore, collocations within 200 km of land are excluded as these are typically in regions with highly variable (both temporally and spatially) currents. Using this collocated dataset, daily bias and error statistics are calculated in NRT for both TMI and AMSR-E.

Two known problems with the TMI radiometer and retrieval algorithm increase the bias for certain retrievals. As detailed in Gentemann (accepted JGR) there is a bias in TMI data due to an incomplete correction the effects of an oxidation of the TMI antenna coating that causes an error dependent on the temperature of the antenna. A correction was developed but a small residual error that is a function of local time remains. This will be calculated and added to the mean global daily bias. In addition, TMI SSTs have a slightly different dependence on SST and wind speed than AMSR-E SSTs. This is easily corrected through TMI and AMSR-E collocations. The bias, TMI minus AMSR-E is calculated as a function of SST and wind speed. The mean bias is also calculated. To correct TMI for this slight difference, the bias as a function of SST and wind speed minus the mean bias is subtracted from TMI data is added to the mean global daily bias.

The sd. at each pixel is equal to whichever is larger, the daily NRT buoy sd. or the sd. calculated as a function of scene. SST as a function of scene is equal to the root mean square sum of the sensor NEDT, emissivity uncertainty, side lobe contamination, and spatial resolution. The increase in measurement error due to uncertainty in the emissivity is different for various environmental scenes and different for the each radiometer. TMI retrieves SST primarily from the 10.7 GHz channel which has decreased sensitivity (and therefore increased error) at SSTs below 285 K. AMSR-E has a 6.9 GHz channel that accurately retrieves SST at temperature as low as -1.8 C. Both sensors have increased error at high wind speeds (above 12 m/s). These errors should not vary temporally or spatially since they are due to uncertainty in the estimate of emissivity. They are best derived through global comparisons to Reynolds SST. Since we wish to examine the variability of the sd. in AMSR-E, we assume that the minimum sd. as a function of SST and wind speed is due to the difference between Reynolds and AMSR-E. Therefore, we can calculate the increase in sd. as a function of SST and wind by calculating sd. (SST,Wind)-minimum(sd.(SST,wind)). TMI's equatorial orbit yields relatively few collocations with SST below 285K and wind speeds above 12 m/s, therefore the TMI algorithm errors are better calculated from the AMSR-E radiometer measurements run with the TMI algorithm.

Finally, since the MW SST data is at a lower resolution than the final GHRSST-PP product (1/12°), the sd. can be calculated as a function of geographic location based on a high resolution SST product (such as ATSR or possibly the 1/16° Navy model SST) averaged down to the resolution of the TMI and AMSR-E radiometers. The sd. at MW resolutions minus the sd. at 10 km (~1/12°) should be included as part of the error estimate. This calculation should include several years of data, from ENSO and 'normal' years. The largest sd. difference for a given location should be used. Although it is very likely that the sd. due to differing spatial resolutions will change in time, taking the largest sd. at any given location should encompass most situations. Another methodology that may be investigated is to utilize the previous month of L4 analysed SST products to calculate how the lower spatial resolution varies geographically. This will have a temporal lag, but will perhaps vary more accurately with differing oceanic modes of variability. There should be some error due to the emissive antenna on TMI and the hot load calibration correction for AMSR-E, but it is unclear how to include a correction for these effects at this time.

In order to simplify the process of SSES derivation for both TMI and AMSR-E data, reference tables for calculation of bias and standard deviation as a function of geolocation, local time, proximity to error sources, SST, and wind speed will be delivered to the GHRSST-PP by RSS. Initial reference tables will be made available by February 2004.

WP-ID3.3 Publish SSES on GHRSST-PP web site and inform RDAC/GDAC

SSES will be computed by the GHRSST-PP Science Team and published at regular intervals on the GHRSST-PP User Information Services Web portal. In addition, SSES will be sent to the main point of contact for each GHRSST-PP RDAC project.

It is expected that the derivation of SSES will evolve as the GHRSST-PP science Team becomes more experienced with their derivation and application. SSES data will be maintained by the GHRSST-PP International Project Office and the GHRSST-PP Science Team.

WP-ID4: Generation of L4 Analysed data products (L4)

WP-ID4	Generation of Analysed data products for each processing window
Work Package number :	**WP-ID4**
Leader:	**GHRSST-PP Science Team, RDAC project leaders, GHRSST-PO**
Aim:	**To specify the data format, processing method and algorithms that will be used to generate L4 analysed data products (L4)**

1

Objectives:

1. Specification of the data processing method and algorithms that will be used to generate L4 SSTfnd (L4SSTfnd) data products
2. Specification of the L4SSTfnd MMR_FR metadata that should be prepared and registered at the MMR
3. Specification of HR-DDS data granules and associated HR-DDS MMR metadata that should be extracted from each L4SSTfnd data product

2

Description:

The GDS will produce an analysed estimate of SSTfnd every 24 hours using analysis procedures that depend on L2P data products as input. GDS L4SSTfnd data products provide complete coverage SST data fields that are free of gaps for each APPW.

For the first version of the GDS, the GHRSST-PP Science Team have specified that a diversity of approach is required to establish the optimal analysis method that will maximise the impact of complementary satellite observations. Standard GHRSST-PP L4SSTfnd products will be generated on a grid scale of 1/12° every 24 hours. Ultra-high resolution (UHR) L4SSTfndUHR data products will be generated for regional areas at a grid scale of 1/48° latitude x longitude (~2km) every 24 hours.

This work package specifies the "rules" and procedures that will be used to generate L4 data products, register each product at the MMR and, extract HR-DDS data granules and associated metadata.

3

Inputs:

1. L2P data files
2. L4 data processor configuration file

4

Outputs:

1. Global coverage L4SSTfnd data product (L4FND) and associated confidence and error statistics for each APPW
2. Regional coverage L4SSTfndUHR data product (L4FNDUHR) and associated confidence and error statistics for each APPW
3. MMR_FR metadata record for each L4 product
4. Extracted HR-DDS data granules and HR-DDS MMR_FR metadata record for each data granule

5

Acceptance tests:

1. Identical L4 output data products are produced at each RDAC and GDAC when the GDS test data set us used compared to the reference processor.
2. L4 MMR_FR metadata entries are timely and accurate.
3. HR-DDS L4 data products are extracted correctly
4. L4 HR-DDS-MMR –FR metadata is correct and accepted by the HR-DDS MMR system

6	**Metric for performance assessment:** 1. L4 products and HR-DDS L4 granules are produced in a timely manner 2. L4 MMR_FR metadata records provided and ingested at MMR and HR-DDSS MMR in real time. MMR and HR-DDS MMR rejection rate < 2% 3. OPLOG L4_DPSR records correct and timely

This part of the GDS describes the rules and methods used to generate L4 analysed data products. L4 SST data products provide an estimate of the foundation SST (SSTfnd). At 24 hour intervals, called Analysed Product Processing Windows (APPW) defined in Table 2.3.1, global coverage L2P data products are used in an analysis procedure. Two L4 data products are specified by the GDS: an estimate of SSTfnd at a grid resolution of 1/12° having global coverage and an Ultra High Resolution SSTfnd product at a grid resolution of 1/48° having regional coverage. Each data product will include an estimate of the diurnal variability based on a parameterisation that can be driven using ancillary solar radiation and wind speed data and will be formatted to a common GDS L4 data format (described in Appendix A1). These data products are based on user requests that are described in the GDIP and will be made available to the user community in real time.

Figure 4.1 presents a functional breakdown diagram of the data processing steps required to produce L4 data products for each APPW. The processor takes L2P data files as input together with a L4 configuration file containing threshold and data processor configuration settings. For each APPW L2P data are first collated for each APPW and grid cell. L4SSTfnd is calculated using analysis procedures for each output grid cell (WP-ID4.2 in Figure 4.1). An estimate of the diurnal variability within a 24 hour period is then derived for each grid cell (WP-ID4.3 in Figure 4.1). Each L4 data product is then formatted as a L4 data file (WP-ID4.4 in Figure 4.1) and a separate MMR_FR is prepared and submitted to the MMR (WP-ID4.5 in Figure 4.1). HR-DDS granules are extracted from each L4 data product (WP-ID4.6 in Figure 4.1) and a HR-DDS MMR_FR prepared and submitted to the HR-DDS MMR system (WP-ID4.7).

There is currently no consensus within the GHRSST-PP science Team or indeed the international SST community that defines the optimal analysis procedure to use for the generation of high resolution SST data products based on multiple satellite data inputs. There are several schemes that have been developed by operational agencies and research establishments (e.g., Reynolds and Smith) that could be used but in each case, the specific method must be adapted to the high resolution space scales that the GHRSST-PP is attempting to resolve. Significant developments have already produced viable results at the Japanese RDAC through the New Generation (NGSST) SST project and further developments are expected from other approaches at the EU RDAC (within the Medspiration project) and USA RDAC/GDAC (within the NOPP SST for GODAE project) in the next 12 months.

However, these groups first require L2P data files to work with before analysis procedures can be operationally implemented within the GHRSST-PP. Consequently, it is expected that L4 data products will evolve as RDAC gain more experience with high resolution SST measurements. In particular, much is to be learned based on the results obtained from analysis of independent SSES and HR-DDS granules together with in situ observations.

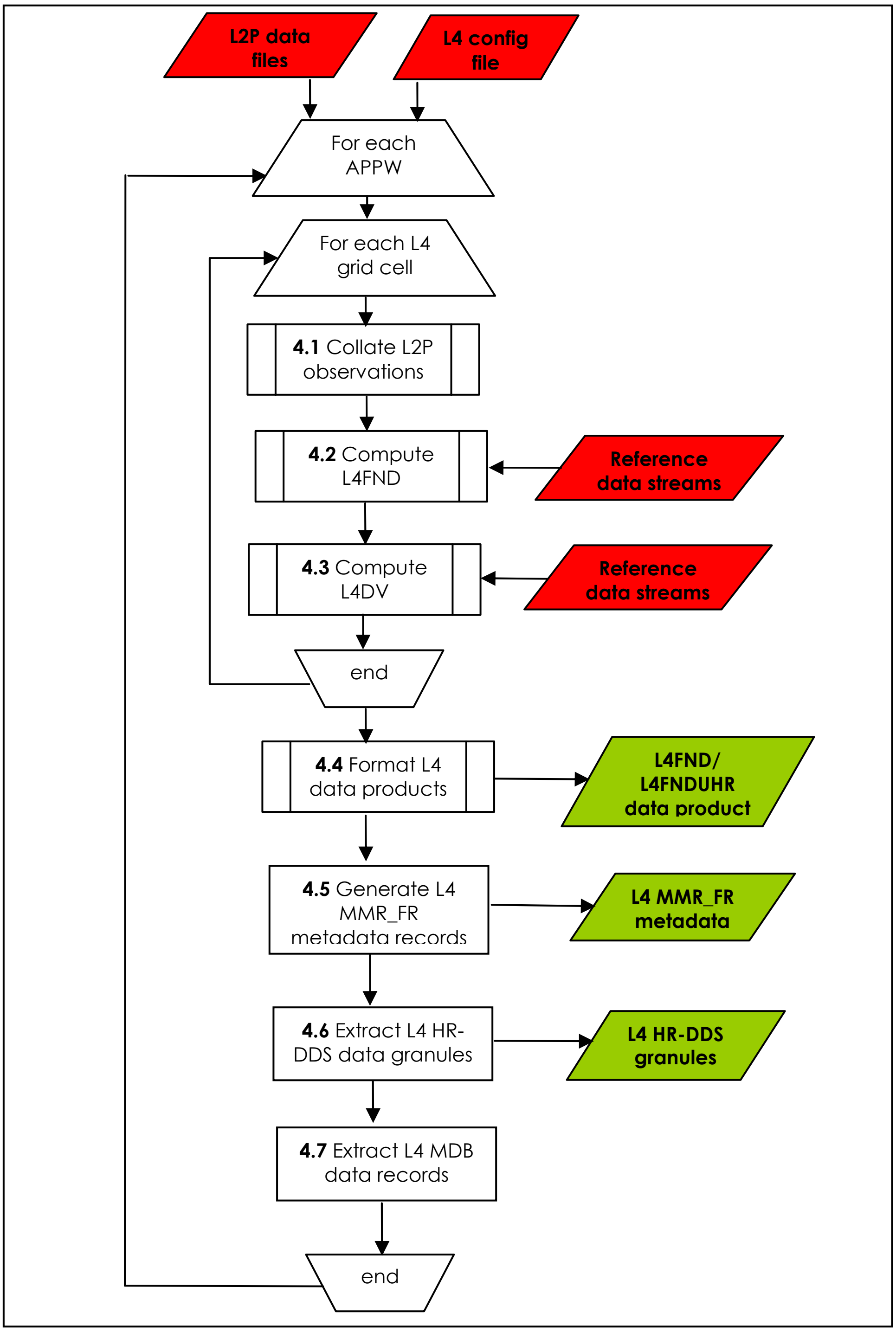


Figure 4.1 Functional breakdown diagram for WP-ID4 showing the data processing steps required to generate GDS L4 SST data products.

The following sections specify the basic approach to developing SSTfnd data products and in particular, the boundary conditions that RDAC and GDAC must consider when developing analysis systems.

WP-ID4.1 Collation of Available L2P Data

Before any analysis can be performed, all L2P data eligible for analysis must first be collated. Some analysis schemes will make use of all available data sources within the analysis procedure whereas others will select the most useful data that are available and may reject all others (through simple ranking or by computing cross-correlations for example). L2P data must then be mapped into the analysis grid.

L2P data must first be re-sampled onto the L4FND standard output grid which is defined in Appendix A1.1. A number of input sensors (infrared imagers) have finer spatial resolution than the output grid specification, as shown in Figure 4.1.1. For these cases the rule will be to average the values of all pixels which overlap the product cell entirely and which have a L2P confidence record **`proximity_confidence`** value equal to the highest encountered within the cell, to produce a single value.

The following rules are specified by the GDS:

- **Rule 4.1a:** In the case of a smaller L2P input pixel than the L2FND grid cell size, L4FND data product cell values are derived from an average of the L2P pixel which completely overlap the product cell and which have a L2P confidence record **`proximity_confidence`** value equal to the highest encountered within the cell, to produce a single value.
- **Rule 4.1b:** For input pixels that straddle the boundary between output grid cells, a weighting function may be applied to the input values according to the degree of coverage of the output grid cell and according to the SSES.
- **Rule 4.1c:** The SSES value for a re-gridded data set will take the value associated with the **`proximity_confidence`** value assigned by Rule 4.1a.

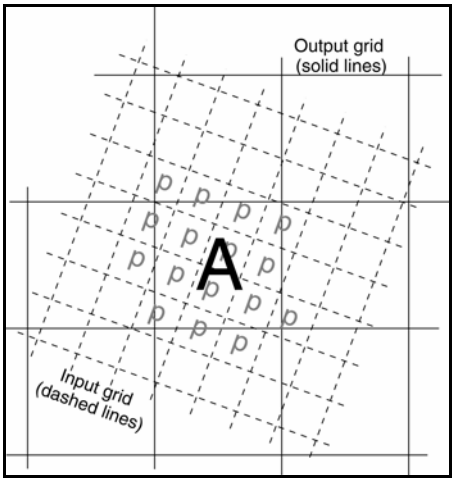


Figure 4.1.1 To illustrate the approach when the L4FND product output grid is over-sampled by the L2P input data. All pixels labelled *p* in the input data are possible contributors to the value for new cell *A*.

When pixels in the input data stream are much larger than those of the SSTFND output grid (e.g., microwave instruments), as illustrated in Figure 4.1.2, the output grid cell takes the value and confidence data record of the input pixel in which it lies. Should it straddle more than one input pixel, it will take a weighted average of all those input pixels with which it overlaps and which have the highest L2P confidence record **`proximity_confidence`** value amongst the overlapping pixels. Weights will be assigned based on the coverage an input pixel has over the output grid.

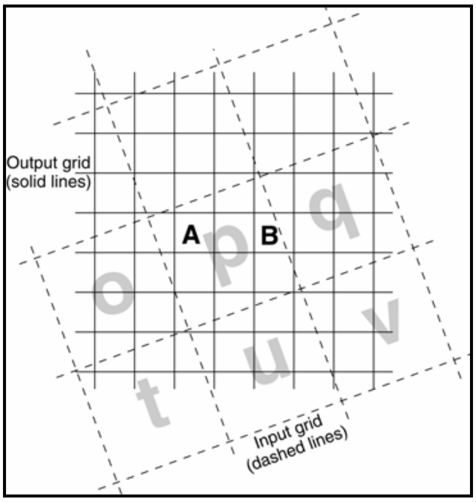


Figure 4.1.2 To illustrate the approach when the L4FND output grid is under-sampled by the L2P data. Grid cell *A* is assigned the value of pixel *p*. Grid cell *B* is assigned the weighted average of p and q provided they both have quality flags with the same rating.

The following rules are specified by the GDS:

- **Rule 4.1d:** In the case of a larger pixel than the L4FND grid cell size, the output grid cell takes the value of the L2P pixel in which it lies.
- **Rule 4.1e:** In the case where the output grid cell is covered by more than one L2P pixel, it will take the weighted average of all those L2P pixels with which it overlaps and which have the highest L2P confidence record `proximity_confidence` value amongst the overlapping pixels. Weights will be assigned based on the coverage an input pixel has over the output grid.

There are other optimal methods for re-gridding data that may be investigated by the GHRSST-PP during the implementation of the GDS-v1.0 and may provide improvements over the simplistic method described here. Furthermore, in most optimal analysis systems, the analysis itself will optimise the re-gridding of data although re-gridding prior to interpolation provides a more straightforward processing system.

WP-ID4.2 Generate L4 SSTfnd data products (L4SSTfnd)

Several methods are available for the generation of complete SST fields from incomplete L2P input data (e.g., Reynolds and Smith, 1994; Reynolds et al, 2002; Guan and Kawamura, 2003; Murray et al. 2002; Murray et al. 1994; Fieguth et al., 1998;2000, Menemenelis et al. 1997; Lorenc, 1981 none of these methods has been fully developed for use in an operational high resolution processor such as the GDS demands. Bias correction of all input data to the analysis procedure is critical to obtaining a valid output (see for example Reynolds et al, 2002). The GDS will use the L2P SSES derived bias as a measure of the overall uncertainty associated with each input data stream. In addition, bias due to diurnal stratification and cool skin effects must also be accounted for using additional data. Some satellite sensors provide a direct estimate of the SSTskin (AATSR) that must be adjusted to the SSTfnd. These data can then be used in an OI procedure to derive complete global fields. As the GHRSST-PP Science Team and RDAC project teams gain experience, it is expected that the bias correction strategy will evolve and an upgrade of this component of the GDS is expected.

L4 UHSSTfnd data products will be generated for regions shown in Table 4.2.1. Other areas are expected in the future (e.g., European coastal seas) and will be described in Table 4.2.1 when operations begin.

Table 4.2.1 L4 UHSSTfnd products that are now in development within the GDS-v1.0

Name	Regional project	Area
Mediterranean	Medspiration (EU-RDAC)	30°N to 46°N and from 6°W to 36.5°E at 2 km spatial resolution
NEAR-GOOS region	NGSST (Japan RDAC)	North-east Asian region bounded by China, Republic of Korea, Democratic Republic of Korea and the Russian Federation along the western boundary and by the Russian Federation and Japan along the eastern boundary. It is part of the WESTPAC region.

The GDS analysis scheme used to generate L4FND data products must:

(a) Provide a daily global coverage combined SSTfnd product that builds on the synergy between complementary SST data streams having a grid cell size of 1/12° latitude x longitude (GDS-v1.0),
(b) Provide error estimates for each analysed grid cell,
(c) Account for differences in spatial and temporal sampling characteristics of each data stream,
(d) Account for gaps in coverage due to the presence of cloud, rain or lack of input L2P data,
(e) Account for SST diurnal variability both in space and time,
(f) Provide a measure of diurnal variability within the data product time domain to accompany the SSTfnd estimate that can also provide a best estimate of the skin temperature of the ocean (SSTskin) for each grid cell.

As an initial reference baseline, the analysis scheme described by Reynolds et al. (2002) may be modified to cater for high resolution spatial scales of SST. Appropriate spatial de-correlation length scales must be derived by the GHRSST-PP RDAC projects and close interaction with the GHRSST-PP Science Team.

All parameters associated with the operation of the L4SSTfnd analysis scheme should be stored in a L4SSTfnd configuration file allowing for their easy editing and adjustment. In parallel, it is expected that alternative methods to the scheme proposed by Reynolds et al (2002) will be explored and compared during the GHRSST-PP within the framework of the GODAE inter-comparison project.

WP-ID4.3 Generate L4 diurnal variation data fields

Diurnal variation data fields form part of each L4 data product and are designed to enable a user to estimate the magnitude and phase of diurnal variability at temporal intervals of 6 hours together with the SSTfnd. The diurnal variation will be referenced to the SSTskin. It is expected that as the GHRSST-PP matures, a more comprehensive strategy will be adopted to generate estimates of SSTskin from multiple sensors.

Following discussions at the 4th GHRSST-PP workshop, the GHRSST-PP Science Team agreed that a parameterisation of diurnal variability should be made available to the user community that allows the phase and magnitude of diurnal variation to be estimated at hourly intervals for each SST grid cell. The Diurnal signal will be referenced to the SSTsub-skin temperature so that warm layer effects are included but an additional parameterisation for cool skin effects will be required to obtain the SSTskin. The initial specification is to use the diurnal variation scheme proposed by Alice Stuart-Menteth (see the GHRSST-PP 4th Workshop report, available from

http://www.ghrsst-pp.org) during the daytime. As an initial specification for night-time, the bias correction strategy described in Donlon et al (2002) may be used to bias adjust SSTsub-skin observations to SSTskin.
The Alice Stuart-Menteth sheme requires the knowledge of several consecutive integrated SSI values at the same time. It should be emphasized that the L2P content in its present definition does not allow applying this scheme. Such a correction should thus use the original auxiliary data. The conversion of UTC integrated SSI provided by the auxiliary data into local solar time integrated values needed by the correction scheme is another problem to take into consideration.

WP-ID4.3.1 A summary of the Stuart-Menteth diurnal SST cycle parameterisation; an initial DV parameterisation for the GHRSST-PP GDS

Current empirical models that estimate the diurnal amplitude or shape use a daily mean wind speed and a daily mean/maximum insolation. In order to estimate the diurnal cycle with more confidence and precision, diurnal fluctuations in the wind and insolation should be considered. Wind is never constant and a sudden wind burst in the afternoon of 9 m/s while the rest of the day is around 3 m/s will bring the daily mean wind speed to 4-5 m/s and therefore a model would underestimate the diurnal amplitude. The same applies to the insolation, in particular if the daily maximum insolation is used. Cloud cover will vary through-out the day and the only clear patch in a day could be at noon and therefore a model would over-estimate the amplitude of the diurnal cycle.

Current modelling efforts have been built upon to derive a more complex model of the diurnal cycle. A new parameterisation has been developed which estimates the shape (phase and amplitude) of the diurnal cycle for 1m and the sub-skin, based only on wind and insolation measurements at several times of the day. The 1m and sub-skin levels are chosen to be compatible with GHRSST-PP. No skin diurnal model is produced as it is assumed that the diurnal heating between the skin and sub-skin is almost identical and therefore the sub-skin model can be applied to all satellite measurements (cool skin model would need to be applied to the sub-skin model to include all skin effects). A brief summary of the parameterisation is given below and a complete summary can be found in Stuart-Menteth et al. (2004). A coded version of the parameterisation is provided in Appendix (A10). A flow diagram summarising the model steps is also provided and it is strongly advised that this is refered to when applying the parameterisation.

The day is divided into four parts (Figure 4.3.1):

1. morning cooling	0:00 - time of SST_{min}
2. morning heating	time of SST_{min} - 12:00
3. afternoon heating/cooling	12:00 - t_{pm}
4. evening cooling	t_{pm} - 24:00

where t_{pm} is set to 17:00, with the exception for certain meteorological conditions (described in more detail below), when t_{pm} is set to 12:00 + 2($t_{max\,1m}$ – 12:00). The day is not separated any further to keep the parameterisation as simple and practical as possible for users. For each segment of the day, a second order polynomial curve is derived to describe the diurnal variation, DV, of temperature over that time period. The diurnal variations are referenced to the time of the minimum SST, t_{min}, which is assumed to represent the foundation temperature, T_{fnd}.

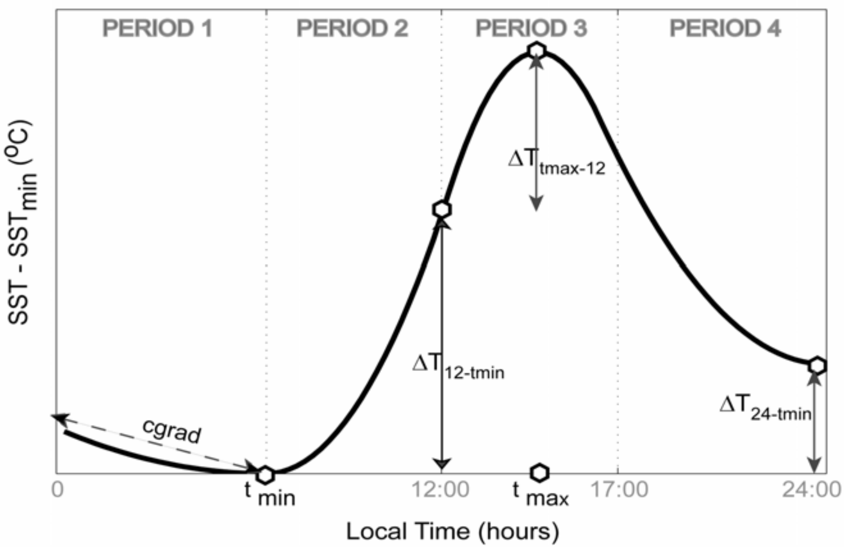


Figure 4.3.1: Schematic of the diurnal SST cycle parameterisation. The day is divided into four parts and six models define the shape of the four curves: cgrad, T_{min}, $\Delta T_{12\text{-}tmin}$, t_{max}, $\Delta T_{tmax\text{-}12}$, $\Delta T_{24\text{-}tmin}$. The predicted diurnal warm layer temperature deviation, ΔT, is referenced from the daily minimum SST where $\Delta T_{tmin} = 0$.

To define the characteristics of the curves and consequently the overall shape, six parameters are required:

(i) Cooling rate in the morning	*cgrad*
(ii) Time of minimum SST	t_{min}
(iii) Morning diurnal heating amplitud	$\Delta T_{(12h\text{-}tmin)}$
(iv) Afternoon diurnal heating amplitude	$\Delta T_{(tmax\text{-}12)}$
(v) Time of daily maximum SST	t_{max}
(vi) Diurnal heating amplitude at midnight	$\Delta T_{(24h\text{-}tmin)}$

The form of the curves is described below:

Period 1 (t = 0:00 – t_{min})

$$DV_t = cgrad.t_{\min}\left(\frac{(t-t_{\min})^2}{(0-t_{\min})^2}\right) \quad \textbf{(Eqn.4.3.1.1)}$$

Period 2 (t = t_{min} – 12:00)

$$DV_t = \Delta T_{(12h-t\min)}\left(\frac{(t-t_{\min})^2}{(12-t_{\min})^2}\right) \quad \textbf{(Eqn.4.3.1.2)}$$

Period 3 (t = 12:00 – t_{pm})

$$DV_t = \Delta T_{(t\max-t\min)} - \left(\frac{(\Delta T_{(t\max-12h)})(t-t_{\max})^2}{(12-t_{\max})^2}\right) \quad \textbf{(Eqn.4.3.1.3)}$$

Period 4 (t = t_{pm} – 24:00)

$$DV_t = \Delta T_{(24h-t\min)} + \left(\frac{(DV_{(t_{pm})} - \Delta T_{(24h-t\min)})(t-24)^2}{(24-t_{pm})^2}\right) \quad \textbf{(Eqn.4.3.1.4)}$$

There two adaptations of the curve for the afternoon heating/cooling period (Period 3) based on extensive examination of many diurnal SST signals.

1. The first case is for the afternoon heating/cooling curve. If the time of the maximum SST at 1m is less than 14h30 then the length of time period 3 is reduced from 12h00 to 17h00 to 12h00 to 12h00+2(*tmax*-12h00). This subsequently alters the start of Period 4 from 17h00 to 12h00+2(*tmax*-12h00). This produces a more realistic cooling shape at 1m during moderate wind conditions.

2. The second adaptation is only for the sub-skin. An additional curve is derived for the afternoon heating/cooling period of the sub-skin to represent cases where there is an abrupt break down in the diurnal stratification. In this situation, the top meter is strongly stratified in the morning so even a small increase in wind can break down the stratification in the top layer and cause the sub-skin SST to drop rapidly. For this case, the new curve computed in addition to Eqn.4.3.1.3.

$$DV_{(t=t_{\max sub}:t_{\max 1m})} = \Delta T_{\max 1m} + \left(\Delta T_{\max sub} - \Delta T_{\max 1m}\right)\left(\frac{\left(t - t_{\max 1m}\right)^2}{\left(t_{\max 1m} - t_{\max sub}\right)^2}\right) \qquad \textbf{Eqn.4.3.1.5}$$

where

$$\Delta T_{\max} = \Delta T_{(12h - t\min)} + \Delta T_{(t\max - 12h)} \qquad \textbf{Eqn.4.3.1.6}$$

An example of this different curve definition is demonstrated in Figure 4.3.2. The first plot shows the results using the principal curve equations for the sub-skin (grey) and 1m (black). The second plot shows the altered curve for the sub-skin to represent the sudden drop in sub-skin SST as water is mixed downwards by an increase in wind speed.

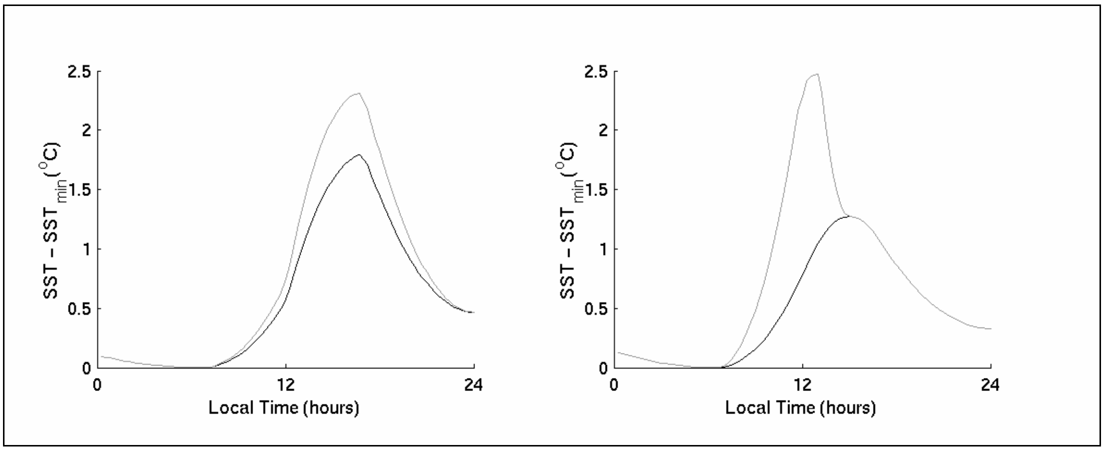


Figure 4.3.2: Example of diurnal cycle parameterisation for 1m (black) and sub-skin (grey) during low wind conditions. In the second plot, wind speed has increased at mid-day and the surface water gets mixed down resulting in a sudden drop in sub-skin SST. A different curve is applied to the afternoon time period to represent this.

Individual empirical models have been derived to compute the six parameters required to define the DV curves, using only mean wind and insolation measurements averaged over our periods of the day (Tables 4.3.1 & 4.3.2). These models were derived using highly accurate buoy measurements from Woods Hole Oceanographic Institution (WHOI) Improved Meteorological (IMET) buoys at several different locations (Arabian Sea, Warm Pool, Sargasso Sea). The models have been validated using other IMET data as well as TAO and PIRATA buoys, providing an independent test data set. There were no sub-skin measurements so these have been derived from the Fairall warm layer model (Fairall et al., 1996). The sub-skin SST measurements often reveal very large ΔT amplitudes of over 4°C. It is difficult to assess the accuracy of the warm layer correction for very low wind conditions as little observational data exists. However, the Fairall warm layer model was compared against some in situ skin data from TOGA COARE, under very low winds, and the predicted warm layer appeared to be good. The results highlight that satellites will measure a much stronger diurnal amplitude than in situ measurements at low winds and supports the need for the diurnal signal to be identified in satellite data.

Table 4.3.1: Description of the variables required to define the shape of the diurnal cycle.

	P1: 0h – t_{min}	P2: t_{min} – 12h	P3:12h – t_{pm}	P4: t_{pm} – 24h
Parameter	*cgrad*	t_{min} $\Delta T_{(12h-tmin)}$	t_{max} $\Delta T_{(tmax-12h)}$	$\Delta T_{(24h-tmin)}$
Variable required	U_{0-6h}	U_{8-12h} $Q_{sw\ 6-12h}$	U_{12-15h} $Q_{sw\ 12-18h}$	U_{16-24h}

Each of the models was derived empirically by regression against the IMET buoy data. In general, the same model and coefficients are applicable for the 1m and sub-skin diurnal cycle parameterisation. It is only under low winds, when the water between the surface and 1m becomes stratified, that different models need to be applied. In these situations, different models are applied to the 1m parameterisation (at very low winds, the 1m response is dampened as less heat is mixed to that depth).

The form of the different models is described in below and the coefficients are reported in Table 4.3.2. Where different models are required for the 1m parameterisation, the threshold conditions are described. These thresholds were determined from the buoy data analysis. The models have a wind speed limit of 0.5m/s. Winds below this limit should be set to 0.5m/s.

$$cgrad = a + be^{(-0.3U_{0-6h})} \qquad \textbf{Eqn.4.3.1.7}$$

$$t_{\min} = a + b\ln(U_{8-12h}) \qquad \textbf{Eqn.4.3.1.8}$$

$$\Delta T_{12h-t\min} = a + be^{(-600/Qsw_{6-12h})U_{8-12h}} + c\left(Qsw_{6-12h}/U_{8-12h}\right) \qquad \textbf{Eqn.4.3.1.9a}$$

<u>If $U_{8\text{-}12h}$ < us, where us=(0.0064$Qsw_{6\text{-}12h}$)+1.2</u>

$$\Delta T_{12h-t\min} = a + be^{(-600/Qsw_{6-12h})us} + c\left(Qsw_{6-12h}/us\right) + \left(\left(d.Qsw_{6-12h}\right)+e\right)\left(us - U_{8-12h}\right) \qquad \textbf{Eqn.4.3.1.9b}$$

$$t_{\max} = a + b\ln(U_{8-12h}) + cQsw_{6-12h} + d\ln(U_{12-15h}) + eQsw_{12-18h} \qquad \textbf{Eqn.4.3.1.10a}$$

<u>If $U_{8\text{-}12h}$ < us, where us=(0.0064$Qsw_{6\text{-}12h}$)+1.2</u>

$$t_{\max} = a' + b'\ln(U_{8-12h}) + c'Qsw_{6-12h} + d'\ln(U_{12-15h}) + e'Qsw_{12-18h} \qquad \textbf{Eqn.4.3.1.10b}$$

$$\Delta T_{t\max-12h} = a + be^{(-600/Qsw_{12-18h})U_{12-15h}} + c\left(Qsw_{12-18h}/U_{12-15h}\right) \qquad \textbf{Eqn.4.3.1.11a}$$

<u>If $U_{12\text{-}15h}$ < 2 m/s:</u>

$$\Delta T_{t\max-12h} = a + be^{2(-600/Qsw_{12-18h})} + c\left(Qsw_{12-18h}/2\right) \qquad \textbf{Eqn.4.3.1.11b}$$

<u>If $U_{8\text{-}12h}$ < 2.5 m/s & ($U_{12\text{-}15h}$-$U_{8\text{-}12h}$)>0.7 m/s:</u>

$$\Delta T_{t\max-12h} = 0.1 \qquad \textbf{Eqn.4.3.1.11c}$$

$$\Delta T_{24h-t\min} = a + b\Delta T_{\max 1m} + cU_{16-24h} \qquad \textbf{Eqn.4.3.1.12}$$

Table 4.3.2: Coefficients for the regression models described in Table 4.3.3. Coefficients in grey represent the different coefficients that are applied to the 1m diurnal model under certain low wind conditions.

Coeff.	*cgrad*	t_{min}	$\Delta T_{12h\text{-}tmin}$	t_{max}	t_{max} 1m (low winds)	$\Delta T_{tmax\text{-}12h}$	$\Delta T_{24h\text{-}tmin}$
a	0.0034	6.4724	0.009	13.0157	15.75	-0.0541	0.1365
b	0.0486	0.1766	6.3071	2.2835	0.8754	2.057	0.209
c			0.0022	-0.002	-0.0027	0.002	-0.017
d			1.8x10-4	-1.9736	-1.9109		
e			0.0114	0.0042	0.0035		

Table 4.3.3: Statistics of derived models against IMET buoy data. Bias = observation – model. Black: 1m model, grey: sub-skin model.

Models	Correlation	Bias ± std. dev.
cgrad	0.34	0 ± 0.014 °C/hr
t_{min}	0.05	0.02 ± 1 hr
$\Delta T_{(12h\ -\ tmin)}$	0.83 0.92	0.002 ± 0.14 °C -0.08 ± 0.37 °C
t_{max}	0.53 0.64	0.1 ± 1hr 0 ± 1hr
$\Delta T_{(tmax\ –\ 12h)}$	0.88 0.87	0.02 ± 0.15 °C 0.04 ± 0.21 °C
$\Delta T_{(24h\ –\ tmin)}$	0.61	0 ± 0.13 °C

Table 4.3.3 summarises the statistics of the model against the data used in the regression. Bias represents observed – model. The 1m results are in black while the sub-skin statistics are shown in grey. Overall, the sub-skin model performs better than the 1m. This is to be expected because at low winds, the relationship between the meteorological conditions and the amount of diurnal stratification breaks down. At the surface, the lower the wind, the greater the warm layer difference. However at 1m, in very low winds, the temperature doesn't continue to increase as wind decreases because very little heat from the top metre is mixed down to that level. In these situations, the heating at 1m is primarily dependent on the amount of radiation absorbed at that depth. The ΔT models have the highest skill. Due to the extreme sensitivity of the diurnal cycle to wind and insolation fluctuations, the phase of the signal responds almost instantaneously to wind variations and therefore to be able to accurately predict the time of the maximum SST, t_{max}, a high resolution time series is required. The model described here provides a best estimate of the phase and it is the first time the phase of the diurnal signal has ever been parameterised. The model does reasonably well and captures the main variations in the timing of the peak of the signal due to changes in wind or insolation conditions. The time of minimum SST, t_{min}, has the poorest correlation. During this period, the changes in SST are very small and so any slight drop in SST will influence the time of the minimum SST. Consequently the observed t_{min} from which the model was derived may not be precise. An alternative method may be to approximate that t_{min} occurs an hour after sunrise. This resulted in a correlation of 0.064 between the observed and predicted t_{min} (where t_{min} is one hour after dawn).

An example of the parameterisation is shown in Figure 4.3.3. The first example shows moderate wind conditions when the diurnal variation (DV) is the same at 1m and sub-skin. The second example shows low winds throughout the daytime where the top meter is stratified until dusk. The third example shows a case with low winds in the morning followed by an increase in wind and a sudden break down in stratification in the afternoon. The parameterisation has been tested and validated using independent surface and 1m SST measurements. The validation results can be found in Stuart-Menteth et al. (2004).

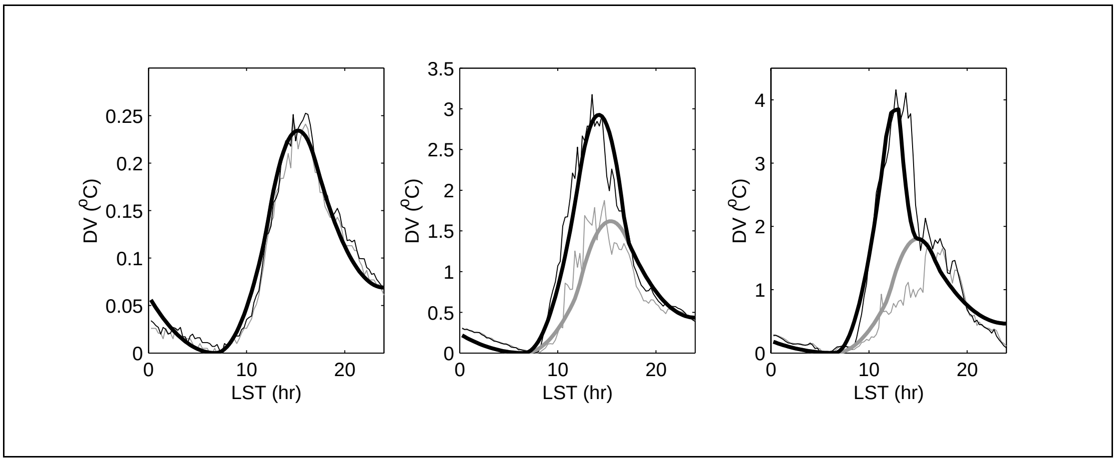


Figure 4.3.3: Example of the parameterisation (DV = diurnal variation from daily minimum SST/$T_{foundation}$). Bold lines represent the parameterisation and thin lines represent observations. Black lines represent sub-skin and grey lines represent 1m.

Caution should be applied when considering removing the DV predicted by the model from the satellite data – especially in periods of very low wind when the shape of the diurnal cycle can change dramatically in immediate response to changes in wind. However, the results shown in Table 4.3.4 suggest that applying the model does reduce the bias. Note that the models have been derived from mean wind speed measurements but single measurements could be used if necessary but obviously limit the accuracy of the model. Table 4.3.4 shows the difference between using a mean and point measurement. The results depend on how variable the wind is at that time of day.

Table 4.3.4: Bias and standard deviation of satellite-measured SST and the foundation temperature with & without the sub-skin model DV corrections. SST(t): satellite SST at overpass time (t), SST_{min}: daily SST minimum (~ T foundation), SST_c: corrected SST (SST-DV), SST_{cu}: corrected SST with point wind measurement at time (t) used for U_{8-12h} & U_{12-15h} instead of a mean value. (Number of data points used, n=270.)

	t = 10h30 (~AMSR/AATSR)	t = 13h30 (~AMSR/AATSR)	t = 14h30 (~AVHRR)
SST(t) - SST_{min}	0.34 ± 0.54	0.68 ± 0.77	0.68 ± 0.62
SST_c(t) - SST_{min}	0.03 ± 0.33	-0.07 ± 0.49	-0.07 ± 0.29
SST_{cu}(t) - SST_{min}	-0.002 ± 0.35	-0.08 ± 0.67	-0.16 ± 0.72

Figure 4.3.4 provides a schematic diagram showing one implementation approach for the Stuart-Menteth scheme within the GHRSST-PP L4 analysis system.

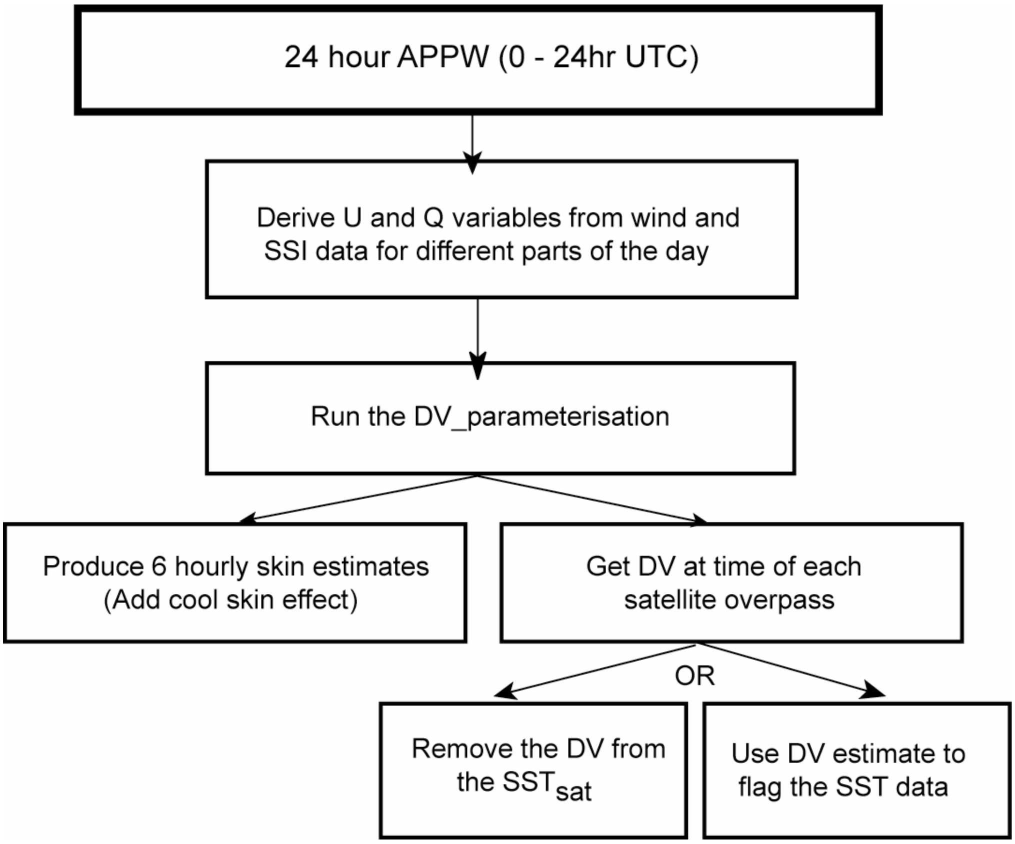


Figure 4.3.4: A suggested way the diurnal warming parameterisation could be implemented into the L4 processing. Parameterisation could be used to remove the diurnal variation (DV) signal at the time of the satellite overpass or to flag the likely hood of diurnal warming.

The DV model requires knowledge of the day's history. If the local solar time of the APPW is 14:00 – 13:59(d+1), then it is difficult to retrieve the morning stratification state, which is required to build the afternoon warming. In this type of situation, the previous APPW information must be retained and made available for computation of the DV.

In Figure 4.3.4, $Q_{sw\ 6\text{-}12}$ & $Q_{sw\ 12\text{-}18}$ must be time average data over the same periods indicated by the subscript. If this is not possible, the mean value could be estimated from an SSI measurement. For example, knowing the time (e.g. t=10am), latitude, longitude, year and year day of a measurement, the value could be estimated, using a common insolation model, as follows:

$SSI_{(6\text{-}12)} = [SSI_{(t)} / Q_{sw\_mod\ (t)}] \cdot Q_{sw\_mod\ (6\text{-}12)}$

where
$SSI_{(6\text{-}12)}$ estimated $Q_{sw\ 6\text{-}12}$
$SSI_{(t)}$ satellite measured SSI at time (t)
$Q_{sw\_mod\ (t)}$ clear sky model Q_{sw} at time (t)
$Q_{sw\_mod\ (6\text{-}12)}$ clear sky modelled mean $Q_{sw\ (6\text{-}12h)}$

It is clear that further work is required with the GHRSST-PP to develop and validate the Stuart-Menteth scheme discussed here.

WP-ID4.4 Format L4 data product

Each L4 data set will be formatted and archived in netCDF format following the data format described in Appendix A1.2, A1.3. The following GDS rules are specified for L4 ancillary data records defined in Table A1.5.2.

- **Rule 4.4.1:** The L4FND grid cell variable **`normalised_analysis_error`** should be updated with the normalised error of the analysis procedure as described in A1.3.3.3.
- **Rule 4.4.2:** The L4FND grid cell variable **`bias`** should be updated with the analysis procedure bias error estimate as described in A1.3.3.4.
- **Rule 4.4.3:** The L4FND grid cell variable **`DT_sst_skin`** should be updated SSTskin estimates computed using a parameterisation/model scheme as specified in A1.3.3.8. The SSTskin parameterisation/model scheme used should be specified in the **`DT_sst_skin`** attribute **`source`** according to Table 4.4.1.

Table 4.4.1 Code values to be used to update the L4FND ancillary data record variable skin_parameterisation_source (Table A1.5.2)

Code	Description
0	No parameterisation specified for this grid cell
1	Stuart-Menteth diurnal variation parameterisation (in press)
2	Donlon et al. (2002) cool skin parameterisation
3	For future use
4	For future use
5	For future use
6	For future use
7	For future use

- **Rule 4.4.4:** If the grid cell is located over land, the L4FND ancillary data record bitfield **`mask`** should be set to 1 as described in A1.3.3.7
- **Rule 4.4.5:** The L4FND grid cell variable **`sea_ice_fraction`** should be set to the most appropriate value derived from the L2P input data streams appropriate to this output grid cell as described in A1.3.3.5. A value of O indicates open ocean and 100 indicates 100% sea ice coverage. In addition, the L4FND grid cell ancillary data field **`sources_of sea_ice_fraction`** should be set according to the source of data used to define the value of **`sea_ice_fraction`** as defined in Table 2.1.1.3.
- **Rule 4.4.6:** The L4FND grid cell variable **`sst_skin_quality_flag`** should be set to provide an indication of the quality and reliability of SSTskin estimates held in the **`DT_sst_skin`** variable as described in A1.3.3.9.

WP-ID4.5 Generate and deliver L4 MMR_FR metadata records

In order for L4 data files to become "visible" to the global GDS system and all other processing centres, a MMR_FR metadata record must be prepared and successfully submitted to the GDS MMR by the responsible processing centre for all L4 data products.

WP-ID4.5.1 Format L4 MMR_FR metadata record

A MMR_FR metadata record should be prepared for each L4 data file according to the specification provided in Appendix A6.2.

WP-ID4.5.2 Register L4 MMR_FR metadata record with MMR system

L4 MMR_FR metadata records should be delivered to the MMR system as soon as possible after the L4 data set has been produced following the procedure described in Appendix A6.4.

WP-ID4.6 Extraction of L4 HR-DDS granules

HR-DDS data granules should be extracted from all L4 data products for each HR-DDS site defined in Appendix A5.5. HR-DDS data granules should be prepared according to the specifications provided in Appendix A5.

WP-ID4.7 Extraction of L4 match up database file records (MDB_FR)

L4* MDB_FR should be extracted according to the procedures described in WP-ID3.

WP-ID5: Validation of GDS data streams

WP-ID5	**Validation of GDS data streams**
Work Package number :	**WP-ID5**
Leader:	**P. Minnett (pminnett@rrsl.rsmas.miami.edu) and the GHRSST-PP Science Team.**
Aim:	**To specify procedures to validate the data streams used within the GDS.**
1	**Objectives:** 1. To specify the procedures used for the validation of GHRSST-PP L2P and L4 data products. 2. To specify the diagnostic tests and tools that will be used to validate the GHRSST-PP methods using HR-DDS data granules and GHRSST-PP data sets. 3. To specify the reporting procedures for validation results to (a) the GHRSST-PP ISDI-TAG (b) the GHRSST-PP community and (c) data providers (as necessary). 4. To specify the feedback mechanisms between the GHRSST-PP GDS and the ISDI-TAG if problems are found and the likely procedure for action (a control loop)
2	**Description:** This section of the GDS describes methods used to validate the data streams within the GDS. Validation of the data products provides the only independent method for assessing data product accuracy and quality and is an integral component of the GDS. WP-ID3 describes the process of matching near contemporaneous independent in situ observations with satellite observations and the derivation of SSES. This WP is dedicated to the analysis of MBD data records to validate GHRSST-PP data L4 data products. The validation of satellite data is a complex process requiring careful attention to QC of all data sets. Consequently, a summary and outline specification for diagnostic tools and procedures is provided as a guide for the on-going validation and assessment of GHRSST-PP data products at RDAC and GDAC. A set of reporting procedures are provided as guidelines to the GHRSST-PP validation effort that are designed to generate appropriate actions and solutions in case of problems as a control loop. In this way decisions relating to changes in the GDS can be properly documented and controlled.
3	**Inputs:** 1. L4 satellite data 2. In situ observations 3. The GHRSST-PP MDB 4. HR-DDS data granules
4	**Outputs:** 1. L4 MDB validation records 2. Diagnostic experiment reports and notifications 3. GHRSST-PP scientific and technical validation reports
5	**Acceptance tests:** 1. L4 MDB data records are ingested by the MDB 2. Diagnostic tools are functional 3. HR-DDS data granules are available and easily accessible

	4. Reporting procedures are functional
6	**Metric for performance assessment:** 1. L4 MDB records are ingested at the MDB with a failure rate of < 2% 2. Appropriate action is taken in case of processing anomalies and adequate feedback is generated

Comparison with independent measurements of SST is the primary mechanism for independently quantifying the accuracy of the satellite SST fields that are used in the GDS. This requires that the input fields and the subsequently merged/analysed fields be compared with high-quality measurements of SSTskin and subsurface SST throughout the period of the GHRSST-PP. The comparisons need not, indeed in some cases cannot, be done in real time, but should be done in as short a time thereafter as practical (in the order of days if possible). This is necessary to ensure that sensor problems are identified promptly and that the consequences of unanticipated geophysical events, such as volcanic eruptions, are quickly identified and their effects on the derived SSTs are quantified without delay. WP-ID3 describes the GHRSST-PP Match Up database (MDB) system is used to store GHRSST-PP validation data records. This work package is dedicated to the analysis of MDB data records.

The objectives of data product validation are several-fold. They include a monitoring of the accuracy of the on-board calibration procedures used by each of the sensors and the specification of regional and seasonal characteristics of the residual uncertainties in the algorithms used to correct for the effects of the intervening atmosphere (clear sky water vapour effects and cloud screening in the infrared; water vapour, surface wind speed and rain-fall effects in the microwave). Validation should also identify errors consequent of the algorithms used to derive the merged and analysed SST fields.

Direct comparison between the separate satellite-derived SST fields is also a useful exercise to assess the relative accuracy of these and to confirm their compatibility before the merging operations. It provides a mechanism for assessing quickly the accuracy of SSTs derived from new sensors that become operational during the GHRSST-PP as a large number of cloud-free comparisons which span the global range of SST and atmospheric variability can be generated in a relatively short time., It also provides a link between observations at high-accuracy but narrow swath (e.g. AATSR) or low resolution (e.g. AIRS) and the independent validating measurements through wide-swath, high resolution sensors (e.g. AVHRR, MODIS) which can act as a transfer standard. Otherwise, achieving the number of coincident measurements between these sensors and the in situ data that are necessary to define the uncertainties through a wide sample of parameter space may take too long.

The individual SST fields that require validation are presented in Appendix A3.2, and the in situ measurements that will be used for the validation include well-calibrated radiometers and interferometers mounted on ships and aircraft, subsurface measurements taken from moored and drifting buoys, the latter including surface drifters and autonomous profilers. These are presented in Appendix A3.4. Auxiliary supporting satellite data streams that can be used to assist in the validation exercise are listed in Appendix A3.3

Over the duration of the GHRSST-PP, the availability of validating sensors is likely to be those that are available now, and which are being used to validate the current generations of satellite sensors. New shipboard and aircraft sensors may become available in the duration of the GHRSST-PP, but confidence in the accuracy of these sensors must first be established before their data are incorporated into the GHRSST-PP. One mechanism for this would be through collocated deployments with established sensors on the same ship or aircraft.

The instruments used to provide the validation data must themselves be subject to accurate and continuing calibration, traceable to national standards, and be deployed according to the appropriate guidelines and protocols.

Because of the different volumes of data involved, it is more efficient for the in situ measurements be provided to the RDAC, instead of the satellite data being delivered to the supplier of the validating measurements. The routinely-collected validation data could be delivered by periodic ftp-pull to the RDACS, while the episodic data could be delivered by ftp-push or by ftp-pull after special notification.

WP-ID5.1 The procedures and rules that will be used to validate GHRSST-PP data products.

The procedures to be adopted for the validation of the GHRSST-PP SST fields are those required to fill the elements in the MDB validation Records (described in Appendix A4.1) and to conduct the comparisons between the satellite-derived SST measurements and the reference data. The information in the validation data sets should be determined and reported in terms of the user requirements. These include not only absolute uncertainties but also relationships and systematic dependences between these errors and other geophysical variables that are represented in the auxiliary data sets.

The following GDS rules are specified:

- **Rule 5.1a:** The validation will be expressed in terms of the statistics of the relationship between the measured satellite SST field and the reference values (measured – reference; i.e. an error, not a correction).
- **Rule 5.1b:** As a minimum the error will be expressed as a mean and standard deviation. But, wherever possible, dependences of the errors on time, geographical location, and the auxiliary variables, such as surface wind, water vapour, aerosols, etc, should be explored and reported.

It is expected that significant scientific study, on a case by case basis, will be required to derive accurate validation reports using MDB data records. Accordingly, GHRSST-PP scientific and technical reports will form the main output of the validation activities.

WP-ID5.2 Diagnostic tests based on the GHRSST-PP high resolution diagnostic data set (HR-DDS)

The GHRSST-PP high resolution diagnostic data set (HR-DDS) provides an extremely rich source of information that can be used to conduct diagnostic experiments on GHRSST-PP data products and L2P data streams. After each satellite product has been calibrated and quality controlled as accurately as possible, several diagnostic products will be computed to assure that the input satellite data are as accurate as possible. This is necessary because the errors change with time, e.g., orbits decay or may be modified due to mission requirements, satellite instruments age and are replaced by new instruments. These diagnostics will be performed on small time and space scales within the HR-DDS system. However, some of the comparisons that are required involve in situ data which are sparse and only helpful on relatively large time and space scales. In addition, it is necessary at this stage to develop a relatively small set of diagnostics so that they are easier to monitor.

WP-ID5.2.1 Additional Tools

In addition to the in situ SSTs produced from drifting and moored buoys, it is planned to use two additional analysis products. The first is the optimum interpolation version 2 (OI.v2) analysis produced by a blend of AVHRR and in situ data (ship and buoy). The analysis is produced operational weekly on a one-degree spatial grid and is described in Reynolds, et al. (2002). The second is the monthly climatology produced from OI.v2 and other analyses with a base period of 1971-2000. The analysis method is described in (Smith and Reynolds, 1998). Xue, et al. (2003) have updated the base period using the Oi.v2 and describe the changes in the base period with time. Both data sets are available in real-time at

http://www.emc.ncep.noaa.gov/research/cmb/sst_analysis/

and

http://www.cpc.noaa.gov/products/predictions/30day/SSTs/sst_clim.html, respectively.

These two products are diagnostic tools only and are not included in any final GHRSST product.

The GHRSST-PP will participate in the GODAE data sharing project as both a contribution to the project itself but also to use the inter-comparison framework to investigate the character of different L4FND analysis schemes. This will be developed once L4FND data streams come on line.

WP-ID5.2.2 Time series diagnostics

Within the GDS, the following time series will be computed and posted to the GHRSST web page under a diagnostic results web page.

1. A weekly time series will be produced of the collocated mean and SD. difference of each satellite product with respect to drifting and moored buoys. This will be done separately for day and night and for 4 regions: global, between 30N and 30S, south of 30S and north of 30N.
2. Because the differences in the above time series are sampled at buoy locations, they are not sampled uniformly. Thus an additional set of weekly time series will be produced of the average SST anomaly for each satellite product and for the Oi.v2 analysis. The anomaly will be computed using the OI.v2 climatology interpolated spatially and temporally as required. This will again be done for the 4 regions, above.
3. To monitor differences on shorter time scales, a daily time series will be produced for each satellite product showing the mean and standard deviation of the anomaly. This will be done separately for day and night and be done for 4 regions listed above.

WP-ID5.2.3 Other diagnostics

These initial diagnostics are produced to identify possible problems. Further diagnostic products will be computed as required. The additional products may include regional time series and mean differences between products.

WP-ID5.2.4 Additional notes on satellite operations

As satellites change and replace or upgraded with time, changes in the quality of the data may be impacted. The GHRSST web page will include notes for users for each satellite product. These notes will include but not be restricted to any change in the instrument or orbit that may impact the data.

WP-ID5.3 The procedures and rules that will be used to validate GHRSST-PP data products external the DDS system.

It may be necessary to include particularly rich or promising validation data sets that lie outside of the DDS system. These can be accommodated by focussed validation exercises to make use of these sources, but operational constraints may require that this be done in a ‘reprocessing’ context. If a new geographically well-defined source of validation measurements becomes available, with a reasonable prospect of continuing operation over a significant time interval, then that location can be included in a revised DDS list. This revision will be approved by the "Problem Resolution Board" (Appendix A8) taking into account the likely benefits and the logistical costs involved.

WP-ID5.4 To specify the reporting procedures for validation results to (a) the GHRSST-PP ISDI-TAG (b) the GHRSST-PP community and (c) data providers (if necessary).

The results of the validation and quality assurance exercises will be made available to all interested parties through a series of web pages that graphically portray the findings. ASCII files of the validation records will be available for ftp-get through the same set of pages. An example such a system, the MODIS (Ocean) Quality Assurance Browse Imagery is accessible at http://jeager.gsfc.nasa.gov/browsetool/ and could serve as a model for the GHRSST-PP scheme after tailoring to meet the requirements specific to GHRSST.

It is anticipated that the GHRSST-PP will be able to use the MQABI code-base (http://mquabi.gsfc.nasa.gov) that allows comprehensive browsing and download of GHRSST-PP HR-DDS granules.

WP-ID5.5 To specify the feedback mechanisms between the GHRSST-PP GDS and the ISDI-TAG if problems are found and the likely procedure for action (a control loop)

The effort invested in the validation exercise is wasted unless the information gained can be used to refine the quality flags and improve the products developed by the GDS. This requires a monitoring activity of all of the fields being validated. A 'Monitor' should be appointed to analyse the content of the validation and QA web pages on a daily basis and generate brief reports on a weekly basis. This will be the responsibility of the GHRSST-PP International Project Office (GHRSST-PO).

Weekly reports will be distributed to the GHRSST-PP Science Team by e-mail and posted on the web pages (possibly with initial password protections for Science Team access only). The web pages should allow users to report problems to the GHRSST-PO who should respond directly to the user and include the contents in the weekly reports. A sub-set of the Science Team, including the Chairman, will be appointed to a "Cal-Val Problem Resolution Board" that should meet by conference call on a monthly basis to resolve problems (by assigning the issue to the most competent Science Team Member as necessary), discuss trends and try to anticipate problems and their solutions. They will also recommend to the Science Team changes to algorithms and processing procedures that would lead to improved products. When an anomalous situation arises, such as a sensor failure or a volcanic eruption, an alarm will be raised by the Monitor by email to the "Cal-Val Problem Resolution Board" which should convene a conference call at the earliest opportunity.

The following GDS rules are specified:

- **Rule 5.5a:** Weekly reports of all cal-val activities should be generated throughout the operational phase of the GHRSST-PP. These should be sent to the GHRSST-PP Science Team.
- **Rule 5.5b:** Monthly conference calls should be convened by the Chairman of the Cal-Val Problem Resolution board (CVPRB).
- **Rule 5.5c:** If a Cal-Val problem is found the CVPRB should convene a dedicated conference call at the earliest possible opportunity in order to establish an appropriate course of action.

WP-ID-6 Commissioning and operating the GDS

WP-ID6	Commissioning and operating the GDS
Work Package number :	**WP-ID6**
Leader:	**RDAC Project leaders and the GHRSST-PO**
Aim	**To specify the acceptance tests and metrics that will be used to commission and monitor the GDS.**
1	**Objectives:** 1. To describe the GDS reference processor. 2. To define the GDS reference data set. 3. To specify the acceptance tests that will be used to commission the GDS. 4. To specify the metrics that will be used to monitor the performance of the GDS. 5. To specify the criteria that will be used to raise a significant processing model failure (SPMF) of the GDS. 6. To specify the criteria that will be used to raise a significant processing model anomaly (SPMA) of the GDS 7. To specify the procedures that will be implemented in the case of a SPMA or SPMF to reinstate data processing as soon as possible.
2	**Description:** The GDS will be commissioned to run at RDAC and GDAC facilities within the GHRSST-PP. It is foreseen as a demonstration system having an operational capability. This WP specifies the criteria that will be used to commission GDS at each RDAC and GDAC. A reference test data set (TDS) and a reference processor (RP) will be used to verify that each RDAC and GDAC as implemented the GDS in an identical manner. This WP also specifies the procedures that should be followed in the case of a significant processing model failure (SPMF) or significant processing model anomaly (SPMA).
3	**Inputs:** 1. The GDS reference document 2. Documentation describing each input data stream
4	**Outputs:** 1. GDS test data set 2. GDS reference processor 3. Commissioning and operating manual
5	**Acceptance tests:** [TBD]
6	**Metric for performance assessment:** [TBD]

This part of the GDS describes how the GDS will be commissioned at each RDAC and GDAC using a reference processor system. The reference processor will be a scientific (as opposed to an operational) implementation of the GDS (typically as IDL/Matalb code). The reference processor will be used together with a test data set (TDS) to verify that RDAC and GDAC implementations of the GDS are capable of providing identical output. In addition, the reference processor can be used to assess the implications of proposed GDS upgrades prior to operational implementation.

WP-ID6.1 The GDS reference data processor

The GDS reference data processor should be a scientific implementation of the GDS. A fourth generation programming language (e.g., IDL or MatLab) should be used to implement the reference processor so that visualisation of data variables is relatively straightforward at any point in the processing system. The reference processor should consider the following aspects:

1. The reference processor should ensure that as many (if not all) of the local data processing steps followed at each RDAC specific to regional data streams are included.
2. The reference processor should be made available to RDAC and GDAC as a reference tool that can be used to troubleshoot regional problems.
3. A master version of the reference processor should be implemented and maintained as such at a GDAC centre.
4. the reference processor should include

    (a) All input SST to L2P SST conversions
    (b) L4 derivations
    (c) An example MMR_FR metadata reference generator
    (d) An example HR DDS granule extractor
    (e) An example HR-DDS MMR_FR metadata generator
    (f) An example SSES generator

It is foreseen that RDAC and GDAC centres will provide example code for the implementation of regionally specific I/O and data analysis modules and that the reference processor code base will be "open source".

The reference processor will be developed by the GHRSST-PP Science Team and RDAC projects.

WP-ID6.2 GDS Test Data Set (TDS)

The GDS is a shared system and must be implemented at several different RDAC/DGDAC. If the GHRSST-PP regional task sharing approach is to be successful and global data products produced from independent data processing centres, it is critical that each RDAC and GDAC produce the same quality and standard of data products. The purpose of a test data set is to implement the GDS in a non-operational environment and provide a common data set to test the implementation of the GDS at each RDAC/GDAC. The GDS test data set (TDS) will be defined and processed by the reference processor to provide a reference output data set. Each RDAC and GDAC may then use the TDS to produce an output data set that can be compared to the reference processor output. Differences between each RDAC/GDAC system can be readily identified by comparing the output and problems solved using the TDS.

The GDS reference data set should include example input data streams from all RDAC and GDAC in the native format expected at RDAC and GDAC. Each input data stream should be extracted for a reference data period which is defined as July-August 2002. A master copy of the GDS reference data set should be made available as a DVD-ROM and distributed to all RDAC and GDAC centres.
The TDS will be available to all participants within the GHRSST-PP. The TDS will include the following:

- **Part 1:** 2 days of SST data files having global coverage from each input satellite data stream considered by the GHRSST-PP as defined in Appendix A3. Both polar orbiting and geostationary observations will be included.

- **Part 2:** 1 week of in situ data straddling the same period as the data sets described in Part 1.
- **Part 3:** An example GHRSST-PP Match-Up Database[15] with entries for each sensor [RD-3].
- **Part 4:** Example GHRSST-PP netCDF L2P, SSTfnd, HR-DDS granule format data files and associated metadata records.

WP-ID6.3 Acceptance tests to commission the GDS at each RDAC and GDAC

In addition to the reference processor, a set of basic functional tests should be defined that can be used to ensure that the regional task sharing processing system defined by RDAC and GDAC interactions and data exchange is functional. The aim of these tests is to ensure that the GDS system is working as an integrated system. Metrics must be agreed and assigned by the GHRSST-PP Science Team, RDAC and GDAC teams according to known priorities and limitations

As the GHRSST-PP is a real time operational system it is imperative that during "operational" phases of the project data are made available and delivered in a timely manner. The criteria for the exchange and delivery between GDAC and RDAC and from data provider to the RDAC have differing priorities that are based on a need to provide sufficient time to produce data products. The following GDS rules are specified:

- **Rule 6.3a:** Each input satellite data stream available for a given processing window should be processed to L2P products as soon as possible after delivery to a processing centre. In any event, L2P should be complete within 3 hours of the end of a PW time period in order to allow sufficient time for the generation and delivery of L4 data products.
- **Rule 6.3b:** All metadata records should be automatically delivered to the MMR immediately following successful data product generation. In any case, metadata should be delivered to the MMR no later than 60 minutes following these events.
- **Rule 6.3c**: All data extracted for use within the HR-DDS system should be ingested into local OPeNDAP systems and made available to the GHRSST-PP community within 3 hours following extraction via a Live Access Server (LAS) style interface. Direct ftp access should also be provided. This is important in order that other RDAC/GDAC systems may take advantage of the ability to monitor the quality of GHRSST-PP data streams in real time.

WP-ID6.4 To specify the criteria that will be used to raise a significant processing model failure (SPMF) or a significant processing model anomaly (SPMA) for the GDS

A significant Processing Model Failure (SPMF) is the term used to refer to the situation where an RDAC or GDAC has failed to provide an expected output data stream within the framework of the GDS. For example, an SPMF will be raised if RDAC data products are delivered too late for inclusion into the appropriate PW or APPW.

A significant processing model anomaly (SPMA) refers to a failure of the GDS that does not result in a SPMF but was of sufficient importance to warrant further investigation.

A list of expected SPMF and SPMA situations should be identified by RDAC and GDAC teams as and when they occur. As a guideline the following criteria should be used to establish potential SPMF:

- What are the significant weaknesses in the regional implementation of the GDS?

[15] TDS MDR records may be simulated in order to provide functionality.

- Where are the likely failures expected to occur?
- What should the action be at RDAC to address these failures should they occur

WP-ID6.5 To specify the procedures that will be implemented in the case of a SPMA or SPMF to reinstate data processing

In the case of a SPMA or a SPMF the GDS TAG Problem resolution Board (PRB) will be notified and take appropriate action as described in Appendix A8.

WP-ID7: Upgrade paths from GDS to GDS-v2

WP-ID7	Upgrade paths from GDS to GDS-v2
Work Package number:	**WP-ID7**
Leader:	**GHRSST-PP International Project Office (Craig Donlon craig.donlon@jrc.it)**
Aim	**To identify the components of the GDS that will be upgraded in the GDS-v2**
1	**Objectives:** 1. To identify and document the components of the GDS that require significant upgrade
2	**Description:** As the GHRSST-PP progresses it is clear that significant upgrades to the GDS will be required as more experience is gained during its implementation. This WP is focussed on establishing which components of the GDS will require upgrade and modification to GDSv2.
3	**Inputs:** 1. GDS specification (this document)
4	**Outputs** 1. Within 6 months of GDS commissioning an assessment of upgrade requirements should be complete and reported 2. Within 9 months of GDS commissioning end, a detailed plan for the upgrade of the GDS system should be available.
5	**Acceptance tests:** N/A
6	**Metric for performance assessment:** N/A

The following sections highlight the major components of the GDSv1 processing system that will be upgraded.

WP-ID7.1 Changes in data inputs

It is foreseen that data streams will change throughout the life of the GHRSST-PP. The GDS must be capable of bringing new data streams in and out of the processing system without affecting other components of the system.

WP-ID7.2 Upgrade of IPCV and MPCV procedures

The GDS procedures for the derivation of IPCV and MPCV values described in GDS WP-ID2.1.1.14 are quite basic. It is clear that significant improvements can be made leading to a better quality of pixel error statistics. It is foreseen that this component of the GDS will be upgraded following further research and development within the GHRSST-PP Diagnostic Data Set.

WP-ID7.3 Changes to L2P quality control procedures

It is foreseen that minor modifications to the procedures used to control the quality of input data and derive L2P data files may be required once significant experience is gained with the GDS.

WP-ID7.4 Changes to analysis procedures and data products

The GDS V1.0 is considered a first step towards an operationally efficient analysis scheme that will deliver estimates of the subsurface SST in real time. It is expected that as the GHRSST-PP project develops significant modifications to this methodology will be made in terms of the bias adjustment strategy.

The methodology outlined for the generation of L4SSTfnd is basic in the GDS in order to have an operational system in place. It is foreseen that a significant upgrade of this scheme will be undertaken in the GDSv2.

WP-ID7.5 Changes to data delivery and data format

As experience is gained in the distribution and application of GDS data products, data product format changes are foreseen based on user requirements and data delivery constraints.

WP-ID7.6 GDS reference processor upgrade

The GDS reference processor system should be used to test all proposed upgrades so that the implications of the proposed changes may be evaluated prior to operational implementation.

WP-ID7.7 Changes to the Diurnal Variation parameterisations

It is expected that the diurnal variation paramterisatio0n scheme of Stuart-Menteth will be upgraded following initial implementation.

WP-ID7.8 Computation of SSES

It is expected that the procedures and analysis techniques used to derive SSES will be upgraded.

References

Brisson, A., S. Eastwood, P. Le Borgne and A. Marsouin, O&SI SAF Sea surface Temperatures: pre-operational results, Proceedings of the 2001 Meteorological Satellite Data User's conference, Antalya, Turkey, 2001.

Carter, E. F. and A. R. Robinson, 1987: Analysis models for the estimation of oceanic fields. J. Atmos. Oceanic. Technol. , 4, 49-74.

Casey, K., et al. the GHRSST-PP Reanalysis Project Strategy, (in prep) available from the GHRSST-PP International Project Office, 2003.

Cavalieri, D. J., J. Crawford, M. Drinkwater, W. J. Emery, D. T. Eppler, L. D. Farmer, M. Goodberlet, R. Jentz, A. Milman, C. Morris, R. Onstott, A. Schweiger, R. Shuchman, K. Steffen, C. T. Swift, C. Wackerman, and R. L. Weaver., NASA sea ice validation program for the DMSP SSM/I: Final report. NASA Technical Memorandum 104559, 1992.

Cavalieri, D. and J. Comiso. Algorithm Theoretical Basis Document for the AMSR-E Sea Ice Algorithm, Revised December 1. Landover, MD, USA: Goddard Space Flight Center, 2000. http://nsidc.org/data/docs/daac/ae_si12_12km_tb_sea_ice_and_snow.gd.html and http://polynya.gsfc.nasa.gov/seaice_projects.html.

Donlon, C. J., T. J. Nightingale, T. Sheasby, J. Turner, I. S. Robinson and W. J .Emery, The Thermal Skin Temperature of the Ocean at High Wind Speed, Geophys. Res. Lett., Vol. 26, No. 16, 2505-2508, 1999.

Donlon, C. J., P. Minnett, C. Gentemann, T. J. Nightingale, I. J. Barton, B. Ward and, J. Murray, Towards Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research, J. Climate, Vol. 15, No. 4, 353-369, 2002.

Donlon,. C. J., The GODAE High Resolution Sea Surface Temperature Pilot Project: Strategy and Initial Implementation Plan, available from the GHRSST-PP International Project Office, 2002a.

Donlon, C. J., Proceedings of the Second GODAE High Resolution Sea Surface Temperature Pilot Project Workshop, Tokyo, Japan, May 13-15th 2002, available from the GHRSST-PP International Project Office, 2002b.

Donlon, C. J., Proceedings of the Third GODAE High Resolution Sea Surface Temperature Pilot Project Workshop, ESA/ESRIN, Frascati, Italy, December 2-4, 2002, available from the GHRSST-PP International Project Office, 2003a.

Donlon, C. J., The GODAE High Resolution Sea Surface Temperature Pilot Project Development and Implementation Plan (GDIP), available from the GHRSST-PP International Project Office, 2003b.

Emery, W. J., K. Cherkaur, B. Shannon and R.W. Reynolds, Hull mounted sea surface temperatures from ships of opportunity, J. Atmos. Oceanic, tech., 14(5), 1237-1251, 1997.

Fieguth P.W., Allen M.R., Murray M.J., Hierarchical Methods for Global-Scale Estimation Problems, Canadian Conf. on Elec. Eng., Waterloo, Canada,161-164, 1998.

Fieguth, Khellah, Murray and Allen, Data fusion of SST data, Proc. IEEE 2000, 5, 2111-13, 2000.

Faugere,Y., P.Le Borgne and H.Roquet, Realisation d'une climatologie mondiale de la temperature de surface de la mer a echelle fine, La Meteorologie, 35, 24-35, 2001.

Gardashov, R. and M. C. Barla, the calculation of the distribution of the Sun glitter radiance on the ocean surface by observing from a geostationary orbit, Imt. J. Rem. Sens., 22, No. 15, 2939-2952, 2001.

Gentemann, C. L., C. J. Donlon, A. Stuart-Menteth and F. J. Wentz, Diurnal signals in satellite sea surface temperature measurements, in press, Geophysical Research Letters, 2003.

Guan, L. and H. Kawamura (2002): Merging satellite infrared and microwave SSTs—Methodology and Evaluation of the new SST, Submitted to J. Oceanogr. (Under review)

Guan Lei, H. Kawamura (2003): SST availabilities of satellite infrared and microwave measurements, J. Oceanogr., 59, 201-209.

Hosoda, K. and H. Kawamura (2002a): Examination of the merged sea surface temperature using wavelet analysis, submitted to J. Oceanogr. (Under review)

Hosoda, K. and H. Kawamura (2002b): Time-dependent spatial decorrelation scales of sea surface temperature field derived from the merged sea surface temperature, submitted to J. Oceanogr. (Under review)

Josey, S. A., E. C. Kent and P. K. Taylor, The Southampton Oceanography Centre (SOC) Ocean - Atmosphere Heat, Momentum and Freshwater Flux Atlas. Southampton Oceanography Centre Report No. 6, 30 pp. plus figs, 1998.

Kawai, Y. and H. Kawamura, Evaluation of the Diurnal Warming of Sea Surface Temperature Using Satellite-Derived Marine Meteorological Data, Journal of Oceanography, Vol. 58 (No. 6), pp. 805-814, 2002

Kawai, Y. and H. Kawamura (2002): Evaluation of the diurnal warming of sea surface temperature using satellite-derived meteorological data, J. Oceanography, 58, 805-814.

Lorenc, A., A global three-dimensional multivariate statistical analysis system, Mon. Wea. Rev., 38A, 137-161, 1981.

Menemenlis, D., P. Fieguth, C. Wunch and A. Willsky, Adaptation of a fast optimal interpolation algorithm to the mapping of oceanographic data, J. Geophys. Res.,102, 10,573-10,584, 1997.

Murray M.J., Assimilation Techniques for blending GOES & POES SST retrievals., RALTechnical Report RAL-TR-2002-017, 2002.

Murray, M.J., Allen, M.R., Mutlow, C.T., and Llewellyn-Jones, D.T. Investigating tropical intraseasonal variability in sea surface temperature using data from the Along Track Scanning Radiometer. Proc. of the 2nd ERS-1 Symposium, Hamburg 1993, ESA SP-361, 529-533, 1994

Pinnock S. and C. J Donlon, Implementation of the GHRSST-PP High Resolution Diagnostic Data Set, GHRSST-PP Document reference GHRSST/14, available from the GHRSST-PP International Project Office, 2003.

Podesta, G and Evans, R, AVHRR Pathfinder Oceans Matchup database 1985-1997 (version 19.0), see http://www.rsmas.miami.edu/groups/rrsl/pathfinder/Matchups/match_index.html, 2003.

Reynolds, R. W. and H. Zhang, Buoy need for Sea Surface Temperature (SST), presentation at the 3rd GHRTSST-PP workshop, ESA/ESRIN, Frascati, Italy, 2-4th December 2002, GHRSST-PP Report Number 16, available from the GHRSST-PP Project Office, 2003

Reynolds, R. W., A real-time global sea surface temperature analysis, J. Climate, 1, 75-86, 1988

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes and W. Wang, An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609-1625, 2002.

Rossow, W. and L. Gardner, Selection of a map grid for data analysis and archival, Journal of Climate and Applied Meteorology, 23, 1253-1257, 1984.

Smith, T. M. and R. W. Reynolds, 1998: A high resolution global sea surface temperature climatology for the 1961-90 base period. J. Climate, 11, 3320-3323.

Tanahashi, S., H. Kawamura et al. (2002): Diurnal Variations of Sea Surface Temperature over the Wide-ranging Ocean Using VISSR on board GMS, J.Geophys. Res., (in press)

Wick, G. A. The First Report of the GHRSST-PP In situ and Satellite Data Integration-Technical Advisory Group (ISDI-TAG), GHRSST-PP Document reference GHRSST/13 v1.0, available from the GHRSST-PP International Project Office, 2002.

Xue, Y., T, M. Smith, and R. W. Reynolds, 2003: Interdecadal Changes of 30-Year SST Normals during 1871-2000. J. Climate, 12 (in press).

Appendix A1. GDS data product format specification

The GHRSST-PP data files themselves have been chosen to follow the Climate and Forecast netCDF conventions because these conventions provide a practical standard for storing oceanographic data, and have already been adopted for the Data Sharing Pilot project within GODAE. The netCDF data format is extremely flexible, self describing and has been adopted as a de-facto standard for many operational oceanography systems.

This appendix provides a detailed technical description of GDS output data products. In its initial configuration, the GDS will output four SST data products according to the specifications laid out in Table A1.

Table A1.1 Data product families produced by the GDS.

	L2 Pre-Processed	Analysed SST
Acronym	L2P	L4
Description	Native SST data streams reformatted into netCDF	L4 analysed data provide an estimate of the SSTfnd that is free of diurnal variability.
Grid specification	Native to SST data format	Sections 2.4.7.1 and 2.4.7.2
Temporal resolution	Native to SST data stream	Analysed product processing window See section 2.3
Delivery timescale	As available	Within 12 hours of an APPW (T+12) - see section 2.3
Target accuracy	Native to data stream	< 0.4 K absolute) 0.1 K relative
Error statistics	Native to data stream if available, Sensor Specific error statistics otherwise	sd. and bias for each output grid point (no input data statistics are retained)
Coverage	Native to data stream	Global
Data Content	Table A1.4.2 and A1.4.3	Table A1.5.3
Nominal data format	netCDF see Section A1.2	netCDF see Section A1.2

A1.1 GDS netCDF common data product file structure

The GDS netCDF file format is designed to be compatible with the GHRSST-PP Diagnostic Data Set (DDS) and the GODAE Data and Product Server Project. GDS netCDF files are based on the attribute data tags defined by the Cooperative Ocean/Atmosphere Research Data Service (COARDS) and climate and forecast (CF) metadata conventions. The CF convention generalises and extends the COARDS convention but relaxes the COARDS constraints on dimension order and specifies methods for reducing the size of datasets. The purpose of the CF conventions is to require conforming datasets to contain sufficient metadata that they are self-describing in the sense that each variable in the file has an associated description of what it represents, including physical units if appropriate, and that each value can be located in space (relative to earth-based coordinates) and time (NetCDF Climate and Forecast (CF) Metadata Conventions, Version 1.0-beta5, 2003).

The following web reference documents are applicable to GDS netCDF file formats:

netCDF: http://www.unidata.ucar.edu/packages/netcdf/
CF conventions: http://www.cgd.ucar.edu/cms/eaton/cf-metadata/CF-working.html
COARDS conventions: http://ferret.wrc.noaa.gov/noaa_coop/coop_cdf_profile.html
UDUNITS: http://www.unidata.ucar.edu/packages/udunits/

GDS data files have a netCDF specification that is configured with four common variables as shown in Figure A1.1.

Dimensions (e.g., m x n)	Mandatory
[n x m] geolocation data (specific to Gridded data set)	Mandatory
[n x m] array of SST data (e.g., SST, wind speed etc)	Mandatory
[n x m] array of confidence flags (specific to variable data set)	Mandatory
Global attributes	Mandatory

Figure A1.1 Schematic representation of an GDS netCDF output file identifying the major file components.

In the context of netCDF, a *variable* refers to data stored in the file as a vector or as a multidimensional array. *Global attributes* are used to hold information which applies to the whole file, such as the data set title. The global attributes shown in the example are the minimum set required for a valid GDS netCDF file. Each individual variable can also have its own attributes, referred as *variable attributes*. The dimensions of each variable must be explicitly declared in the *dimension* section.

In this example, the variable SST consists of a 2-D array [n x m] of sea-surface temperature measurements. In order to save disk space, the 32-bit floating point temperature measurements have been scaled onto 16-bit short integers using the **`add_offset`** and **scale_factor** variable attributes. The floating point CSST values can be recovered using:

SSTfloat = (scale_factor × SSTshort) + add_offset

Also associated with the SST variable are variable attributes describing the units, a longer descriptive version of the variable name, and a fill value. S.I units should be used and described by a character string which is compatible with the Unidata UDUNITS package. The SST variable fill value is used to indicate array elements that do not contain a valid measurement.

As the GHRSST-PP output grid is fixed, variables are included which provide the scales for the latitude and longitude axes of the SST array, these are the so-called *coordinate variables*. No fill value attribute is defined because vector coordinate variables should never have missing values.

A series of confidence data flags are supplied in 2 dimensional [n x m] array confidence variables. These data are specific to each scientific data type (L2P and L4).

Note that it is relatively easy to extend the format of a netCDF data file as another variable data set may be appended to the end of the file without modifying the original file format structure.

A1.1.1 GDS netCDF global attributes

The global attributes shown in blue are mandatory for each GDS netCDF data product file are specified in Table A1.3

Table A1.3 global attribute tags for GDS netCDF data product files (Rev 1: 26/02/2004).

Tag name	Format	Description
Conventions	string	A text string identifying the netCDF convention followed. This attribute should be set to "CF-1.0" to indicate compatibility with the Climate and Forecast (CF)

Tag name	Format	Description
		netCDF convention.
title	string	A descriptive title for the data set
DSD_entry_id	string	the DSD entry related to the data file, e.g. USA-RSS-AMSRE-MW-L2-SST
references	String	References that describe the data or the methods used to produce it. Include here the names of all GDS configuration files that have been used.
institution	string	GHRSST-PP data centre code (see Appendix A2 Table A2. 1) where the data were produced
contact	string	A free text string giving the primary contact for information about the data set
GDS_version_id	String	v1r1.4b (*to be updated*)
netcdf_version_id	String	3.5
creation_date	string,"yyyy-mm-dd"	Date the data file was created (UTC)
product_version	string	Release number of the datafile (File_Version in the MMR-FR, see A6.3.2)
history	string	List of the applications that have modified the original data
platform	string	Satellite identifier
sensor	string	Sensor identifier
spatial_resolution	string	Resolution of the product
start_time	string, "hh:mm:ss UTC"	Time of the first measurement in the data file
stop_time	string, "hh:mm:ss UTC"	Time of the last measurement in the data file
start_date	string,"yyyy-mm-dd UTC"	Start date of the data in universal time coordinated (UTC; ~ Greenwich Mean Time)
stop_date	string,"yyyy-mm-dd UTC"	End date of the data in universal time coordinated (UTC; ~ Greenwich Mean Time).
northernmost_latitude	Float	degrees north, range -90° to +90°
southernmost_latitude	Float	degrees north, range -90° to +90°
easternmost_longitude	Float	degrees east, range -180° to +180°
westernmost_longitude	Float	degrees east, range -180° to +180°
file_quality_index	integer	A code value : 0 : unknown quality 1: excellent (no known problems) 2 : suspect (occasional problems, e.g. after launch) 3 : extremely suspect (frequent problems, e.g.with known satellite problems)
comment	string	Miscellaneous information

A1.1.2 GDS netCDF variable attribute definitions

Table A1.5 describes the variable attributes shall be used. Some may not be relevant for certain variables and reference to the variable requirements (as defined in the CDL description of each variable) should be made to establish which are required. The '**`add_offset`**' and '**`scale_factor`**' variable attributes may vary from one dataset to another, depending on the resolution or the characteristics of the sensor in question. Each RDAC/GDAC is free to adjust these attributes to suit their own requirements, since it does not matter to data reading tools which all have to unpack the data. However, these parameters shall be the same for each files of a given dataset.

Table A1.5 Mandatory variable attribute tags for GDS netCDF data product files (Rev 1: 26/02/2004)

Tag name	Format	Description
_FillValue	Depends on variable type	A value used to indicate array elements containing no valid data. This must be of the same type than the storage (packed) type; should be set as the minimum value for this type.
units	string	Text description of the units, preferably S.I., and must be compatible with the Unidata UDUNITS package. For a given variable (e.g. wind speed), these must be the same for each dataset.
scale_factor	must be expressed in the unpacked data type (for instance float whereas the variable may be stored - or packed - as an array of short)	To be multiplied by the variable to recover the original value
add_offset	must be expressed in the unpacked data type (for instance float whereas the variable may be stored - or packed - as an array of short)	To be added to the variable after multiplying by the scale factor to recover the original value. If only one of scale_factor or add_offset are needed, then both should be included anyway to avoid ambiguity, with scale_factor defaulting to 1.0 and add_offset defaulting to 0.0.
long_name	string	A long version of the variable name
valid_min	Same data type as variable	Minimum valid value for this variable once they are packed (in storage type). The fill value should be outside this valid range.
valid_max	Same data type as variable	Maximum valid value for this variable once they are packed (in storage type). The fill value should be outside this valid range
standard_name	string	A standard and unique description of a physical quantity. For the complete list of standard name strings, see http://www.cgd.ucar.edu/cms/eaton/netcdf/standard_name.html.
source	string	Method of production of the original data or original provider
axis	string	Identifies latitude, longitude, vertical or time axis
positive	string	Indicates the real-world direction of a coordinate variable ("up" or "down"). *E.g.*, a depth or atmospheric pressure coordinate would have this attribute set to "down".
coordinates	string	Identifies auxiliary coordinate variables, such as 2-D lat-lon coordinate arrays, label variables, and alternative coordinate variables.

Tag name	Format	Description
grid_mapping	string	Identifies a variable that defines a grid mapping must be provided if data mapped following a projection. refer to CF convention for standard projection names.
comment	string	Miscellaneous information about the data or the methods used to produce it

A1.1.3 GDS coordinate Variable definitions (Rev 1: 26/02/2004)

netCDF coordinate variables provide scales for the space and time axes for the multidimensional data arrays, and must be included for all dimensions that can be identified as spatio-temporal axes.

1. x (columns) and y (lines) grid dimensions are referred either as 'ni' and 'nj' or as 'lat' and 'lon'. If data are mapped on a regular grid (see L4 product format), lon and lat must be used instead of ni (number of columns) and nj (number of lines). ni and nj must be used only when data are mapped on a non-regular grid (curvilinear coordinates) or following the sensor scanning pattern (scan line, swath).
2. Coordinate vectors are used for data arrays located on orthogonal (but not necessarily regularly spaced) grids, such as a geographic (lat-lon) map projections. The only required attribute is **units**. The elements of a coordinate vector array should be in monotonically increasing or decreasing order. The data type can be any and scaling may be implemented if required.
3. Coordinate arrays are used to geolocate data arrays on non-orthogonal grids, such as images in the original pixel/scan line space, or complicated map projections. Required attributes are **units** and **_FillValue**. Elements of the coordinate array need not be monotonically ordered. The data type can be any and scaling may be implemented if required. **Add_offset** and **scale_factor** have to be adjusted according to the sensor resolution and the product spatial coverage. If the packed values can not stand on a short, float can used instead (multiplying the size of these variables by two)
4. A coordinates variable (= "lon lat"): must be provided if non regular lat/lon grid (map projection or swath data)
5. A grid_mapping (= "projection name"): must be provided if data mapped following a projection. Refer to CF convention for standard projection names.
6. '**time**' is the reference time of the data defined as the youngest L2P data record in this file. It is given to facilitate aggregation of all files of a given dataset along the time axis with such tools as DODS or LAS and also as reference for the deviation of each pixel acquisition time. It is intended to optimize the size of the sst_dtime variable (time associated with each SST pixel) which will store the time difference with this reference time.

A1.1.3.1 Geographic regular latitude/longitude grids (Rev 1: 26/02/2004)

This is the simplest case: on such a projection, only two coordinate variables are requested, which can be stored as vector arrays. The following dimension and coordinate variables shall be used for a regular lat/lon grid. No specific variable attributes are required for other variables (like 'sea_surface_temperature' below).

```
netcdf example {
dimensions:
  lat = 1024 ;
  lon = 1024 ;
  time = 1 ;
variables:
  float lat(lat) ;
```

```
  lat:long_name = "latitude" ;
  lat:units = "degrees_north" ;
 float lon(lon) ;
  lon:long_name = "longitude" ;
  lon:units = "degrees_east" ;
 long time(time) ;
  time:long_name = "reference time of sst file" ;
  time:units = "seconds since 1981-01-01 00:00:00" ;
 short sea_surface_temperature(time, lon, lat) ;
  sea_surface_temperature:long_name = "sea surface temperature" ;
  sea_surface_temperature:units = "kelvin" ;
  sea_surface_temperature:_FillValue = -32768s ;
  sea_surface_temperature:add_offset = 273.15 ;
  sea_surface_temperature:scale_factor = 0.01 ;
  sea_surface_temperature:valid_min = -5000s ;
  sea_surface_temperature:valid_max = 5000s ;
  sea_surface_temperature:source = "EUMETSAT SAF O&SI" ;
```

A1.1.3.2 Non-regular latitude/longitude grids (swath data product file) (Rev 1: 26/02/2004)

In this case where data are gridded following the sensor pattern, no projection can be associated and lat/lon have to be stored in 2-D arrays. Dimensions can not refered as lat/lon anymore since x and y axis of the grid are no more related to the latitude or longitude axis. Each variable must explicitly provide a reference to its coordinate variables ('coordinates' variable attribute). The following dimension and coordinate variables shall be used for a non-regular lat/lon grid (swath product file). The specific variable attribute 'coordinates = "lon lat"' is required for each other variables (like 'sea_surface_temperature' below).

```
netcdf example {
dimensions:
 ni = 1024 ;
 nj = 1024 ;
 time = 1 ;
variables:
 float lat(nj, ni) ;
  lat:long_name = "latitude" ;
  lat:units = "degrees_north" ;
 float lon(nj, ni) ;
  lon:long_name = "longitude" ;
  lon:units = "degrees_east" ;
 long time(time) ;
  time:long_name = "reference time of sst file" ;
  time:units = "seconds since 1981-01-01 00:00:00" ;
 short sea_surface_temperature(time, nj, ni) ;
  sea_surface_temperature:long_name = "sea surface temperature" ;
  sea_surface_temperature:units = "kelvin" ;
  sea_surface_temperature:_FillValue = -32768s ;
  sea_surface_temperature:add_offset = 273.15 ;
  sea_surface_temperature:scale_factor = 0.01 ;
  sea_surface_temperature:valid_min = -5000s ;
  sea_surface_temperature:valid_max = 5000s ;
  sea_surface_temperature:coordinates = "lon lat" ;
  sea_surface_temperature:source = "EUMETSAT SAF O&SI" ;
```

A1.1.3.3 Non-regular latitude/longitude grids (projection) (Rev 1: 26/02/2004)

For gridded data using a specific projection (such as stereographic projection), lat/lon have to be stored in 2-D arrays. Dimensions can not refered as lat/lon anymore since x and y axis of the grid are no more related to the latitude or longitude axis. Each variable must explicitly provide a reference to its coordinate variables ('coordinates' variable attribute) and to the related projection ('grid_mapping' variable attribute) described in a specific variable ('stereographic_polar' in the example below; refer to CF convention for standard names). The following dimension and coordinate variables shall be used for a non-regular lat/lon grid (projection). A specific projection coordinate variable shall be added ('polar_stereographic' here), following the CF-1.0 convention. The specific variable attributes 'coordinates = "lon lat"' and 'grid_mapping = "polar_stereographic"' are required for each other variables (like 'sea_surface_temperature' below).

```
netcdf example {
dimensions:
  ni = 1024 ;
  nj = 1024 ;
  time = 1 ;
variables:
  float lat(nj, ni) ;
    lat:long_name = "latitude" ;
    lat:units = "degrees_north" ;
  float lon(nj, ni) ;
    lon:long_name = "longitude" ;
    lon:units = "degrees_east" ;
  long time(time) ;
    time:long_name = "reference time of sst file" ;
    time:units = "seconds since 1981-01-01 00:00:00" ;
  char polar_stereographic ;
    polar_stereographic:grid_mapping_name = "polar_stereographic" ;
    polar_stereographic:straight_vertical_longitude_from_pole = 0. ;
    polar_stereographic:scale_factor_at_projection_origin = 1.0 ;
    polar_stereographic:false_easting = 0. ;
    polar_stereographic:false_northing = 0. ;
  short sea_surface_temperature(time, nj, ni) ;
    sea_surface_temperature:long_name = "sea surface temperature" ;
    sea_surface_temperature:units = "kelvin" ;
    sea_surface_temperature:_FillValue = -32768s ;
    sea_surface_temperature:add_offset = 273.15 ;
    sea_surface_temperature:scale_factor = 0.01 ;
    sea_surface_temperature:valid_min = -5000s ;
    sea_surface_temperature:valid_max = 5000s ;
    sea_surface_temperature:coordinates = "lon lat" ;
    sea_surface_temperature:grid_mapping = "polar_stereographic" ;
    sea_surface_temperature:source = "EUMETSAT SAF O&SI" ;
```

A1.2 pre-processed (L2P) data products (Rev 1: 26/02/2004)

L2P data products are derived from native SST data products. L2P data products consist of the original SST data values that have been reformatted to a netCDF file format. A L2P confidence data record is provided for each SST pixel, as described in A1.4.2.

A1.2.1 L2P filename convention.

The GDS filename convention used for L2P data products has been designed to provide useful information in an easily readable format. All L2P data product filenames are derived according to the following convention:

<Date Valid>-<L0 ID>-<Processing Centre Code>-L2P-<SST filename>[-<optional characteristic>]-Processing Model ID>.<base format>

which is described in Table A1.2.1

Table A1.2.1 L2P data product filename components

Name	Definition	Description
<Date Valid>	YYYYMMDD	Refers to the date for which this particular data set is valid for.
<L0 ID>	Defined in Appendix A2 Table A2.2	Data set name
<Processing Centre Code>	Defined in Appendix A2 Table A2.1	Processing centre code
<SST filename>	Native to SST filename	Filename of input SST data file as given by data provider
<optional characteristic>	string	Free field to distinguish ambiguous cases (such as ascending/descending pass when contained into a single L2 file)
<processing model ID>	v*nn* (where *nn* is the GDS version number, e.g., 01	Version number of the GDS system used to process the data file
<base format>	nc	Generic file format (nc=netCDF)

The valid date component of the filename forms the first part of the string so that data can be easily sorted by date. For example:

20030621-AVHRR16_L-AUST-L2P-LAC20030621A7SST-v01.nc

Refers to a data set that it is valid for 21st June 2003 (20030621), the source data is AVHRR NOAA 16 LAC (AVHRR16_L) that was generated at the Australian RDAC (AUST), it is a L2P data product (L2P), it is based on an input SST file called LAC20030621A7SST that was generated using the GDS version 1 (v01) and is formatted as a netCDF file (.nc).

A1.2.2 L2P data pixel record format specification

Table A1.2.2 describes the content of a L2P pixel data record that should be created for each SST measurement contained within a L2P file.

Table A1.2.2 GDS L2P SST data record format v3.1 (23/02/04). Mandatory fields are shown in light blue following GHRSST-PP Science Team workshop, Los Angeles USA, September 2003, revised in subsequent discussions to remove inconsistencies within the GDS and for netCDF file format.

Name	Definition	Description	Units
sea_surface_temperature		Pixel SST values use attribute 'sea_surface_temperature:source = "..."' to specify the L2 product source	K
sst_dtime	WP-ID2.1.2	Deviation in seconds from reference time defined as the time of the first L2P data record in this file stored in the netCDF global variable **time**. It is implicit that there will be a 0 value and **sst_dtime** is always positive.	seconds
SSES_bias_error	WP-ID2.1.1.13	SSES Bias error based on confidence flags. ranges from -1 to 1 degree, precision 0.01 degree	K
SSES_standard_deviation error	WP-ID2.1.1.13	SSES standard deviation error based on confidence flags. Ranges from 0 to 1.5 K,	K

		precision 0.01 K	
DT_analysis	WP-ID2.1.1 & WP-ID2.2.1	Deviation from previous day (T-1) SSTfnd analysis. If no analysis is available, the reference mean SST climatology should be used defined as R9 in Table A3.3.1. Precision : 0.1 K	K
surface_solar_irradiance	WP-ID2.1.5	Near contemporaneous SSI value. if 6 hourly irradiance : ranges from 0 to 500 Wm-2, precision 2 Wm-2 if 3 hourly irradiance : ranges from 0 to 1000 Wm-2, precision 4 Wm-2 if unique source (recommended), use attribute **'surface_solar_irradiance**:source = "..."'	Wm-2
wind_speed	WP-ID2.1.4	Wind speed value. ranges from 0 to 50m/s. precision : 1 m/s. if unique source (recommended), use attribute **'wind_speed**:source = "..."'	ms^{-1}
sea_ice_fraction	WP-ID2.1.3	Fractional Sea Ice contamination flag. Ranges from 0 to 1, precision 0.01. if unique source (recommended), use attribute **'sea_ice_fraction**:source = "..."'	Percent
aerosol_optical_depth	WP-ID2.1.6	Aerosol optical depth. if unique source (recommended), use attribute **'aerosol_optical_depth**:source = "..."'	Scaled value.
satellite_zenith_angle	WP-ID2.1.7	-90 to +90 degrees	Degree
rejection_flags	WP-ID2.1.8	Bit field: 0=SST out of range 1=Cosmetic value 2=IR_Cloudy 3=MW_rain 4=MW_ice 5=MW_wind 6=Land 7=spare	Flag
sources_of_wind_speed	WP-ID2.1.4	Source of wind_speed value. Source codes must be detailed in the 'ancillary_sources: comment' attribute	Code
source_of_AOD	WP-ID2.1.6	Source of AOD data. Source codes must be detailed in the 'ancillary_sources: comment' attribute	Code
sources_of sea_ice_fraction	WP-ID2.1.3	Source of Fractional sea ice data. Source codes must be detailed in the 'ancillary_sources: comment' attribute	Code
sources_of_ssi	WP-ID2.1.5	Source of SSI data. Source codes must be detailed in the 'ancillary_sources: comment' attribute	Code
wind_speed_dtime_from_sst	WP-ID2.1.4	Time difference of wind_speed measurement from SST measurement	Scaled hours, 25=not known
AOD_dtime_from_sst	WP-ID2.1.6	Time difference of AOD measurement from SST measurement	Scaled hours, 25=not known
ssi_dtime_from_sst	WP-ID2.1.5	Time difference of SSI measurement from SST measurement	Scaled hours, 25=not known
confidence_flag	WP-ID2.2.2	b0 : 1=potential side lobe contamination; b1 : 1=relaxed rain contamination suspected; b2 : 1=TMI SST retrieved in SST < 285K b3 : 1=high wind speed retrieval b4 : 1=sea ice retrieval for MW data b5 : 1= sun glint suspected b6 : 1= L2 native bias and standard deviation;	5 Flags

		b7 : 1= L2 native confidence value " ;	
Proximity_Confidence	WP-ID2.2.1.5 & WP-ID2.2.2.7	Proximity confidence value	Code

A1.2.3 L2P product file Common data format Description Language (CDL)

The ASCII text representation of a netCDF file is typically specified in network Common data Form Description Language (CDL), see

http://www.unidata.ucar.edu/packages/netcdf/guide_12.html

for more information on netCDF and CDL.

A1.2.3.1 sea_surface_temperature variable

The variable 'sea_surface_temperature' will be included with the format requirements shown in Table A1.2.3.1.

Table A1.2.3.1 CDL description of sea_surface_temperature variable

Storage type	Name	Description	Unit
short	sea_surface_temperature	Pixel sst value	K
CDL description			
short sea_surface_temperature(time, nj, ni) ; sea_surface_temperature:long_name = "sea surface temperature" ; sea_surface_temperature:units = "kelvin" ; sea_surface_temperature:_FillValue = -32768s ; sea_surface_temperature:add_offset = 273.15 ; sea_surface_temperature:scale_factor = 0.01 ; sea_surface_temperature:valid_min = -5000s ; sea_surface_temperature:valid_max = 5000s ; sea_surface_temperature:coordinates = "lon lat" ; sea_surface_temperature:grid_mapping = "polar_stereographic" ; sea_surface_temperature:source = " name of L2 source (ex:EUMETSAT SAF O&SI)" ;			
Comments			
The specific variable attribute 'source' shall be used to specify the L2 product source.			

A1.2.3.2 sst_dtime variable

The variable 'sst_dtime' will be included with the format requirements shown in Table A1.2.3.2.

Table A1.2.3.2 CDL description of sst_dtime variable

Storage type	Name	Description	Unit
short	sst_dtime	deviation in seconds from reference time (earliest L2P data record in the data file, stored in the coordinate variable **time**). it is implicit there is a 0 value and sst_time is always positive.	second
CDL description			
short sst_dtime (time, nj, ni) ; sst_dtime:long_name = "time difference from reference time" ; sst_dtime:units = "second" ; sst_dtime:_FillValue = -32768s ; sst_dtime:add_offset = 0 ; sst_dtime:scale_factor = 1 ;			

sst_dtime:valid_min = -32767s ; sst_dtime:valid_max = 32767s ; sst_dtime:coordinates = "lon lat" ; sst_dtime:grid_mapping = "polar_stereographic" ;
Comments
Refer to WP-ID2.1.2 for definition

A1.2.3.3 SSES_bias_error variable

The variable 'SSES_bias_error' will be included with the format requirements shown in Table A1.2.3.3.

Table A1.2.3.3 CDL description of SSES_bias_error variable

Storage type	Name	Description	Unit
byte	SSES_bias_error	SSES Bias error based on confidence flags, ranging from -1 to 1 K, precision 0.01 K.	K
CDL description			
byte SSES_bias_error (time, nj, ni) ; SSES_bias_error:long_name = "SSES bias error based on confidence flags" ; SSES_bias_error:units = "kelvin" ; SSES_bias_error:_FillValue = -128b ; SSES_bias_error:add_offset = 0. ; SSES_bias_error:scale_factor = 0.01 ; SSES_bias_error:valid_min = -127b ; SSES_bias_error:valid_max = 127b ; SSES_bias_error:coordinates = "lon lat" ; SSES_bias_error:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.1.13 for definition			

A1.2.3.4 SSES_standard_deviation_error variable

The variable 'SSES_standard_deviation_error' will be included with the format requirements shown in Table A1.2.3.4.

Table A1.2.3.4 CDL description of SSES_standard_deviation_error variable

Storage type	Name	Description	Unit
byte	SSES_standard_deviation_error	SSES standard deviation error based on confidence flags. Ranges from 0 to 1.5 K, precision 0.01 K.	K
CDL description			
byte SSES_standard_deviation_error (time, nj, ni) ; SSES_standard_deviation_error:long_name = "SSES standard deviation error based on confidence flags" ; SSES_standard_deviation_error:units = "kelvin" ; SSES_standard_deviation_error:_FillValue = -128b ; SSES_standard_deviation_error:add_offset = 100. ; SSES_standard_deviation_error:scale_factor = 0.01 ; SSES_standard_deviation_error:valid_min = -127b ; SSES_standard_deviation_error:valid_max = 127b ; SSES_standard_deviation_error:coordinates = "lon lat" ; SSES_standard_deviation_error:grid_mapping = "polar_stereographic" ;			
Comments			

Refer to WP-ID2.1.1.13 for definition

A1.2.3.5 DT_analysis variable

The variable 'DT_analysis' will be included with the format requirements shown in Table A1.2.3.5. It expresses the deviation from the previous SSTfnd (T-1) value or, if no analysis is available from the mean reference SST climatology (Reference data set R9). Precision : 0.1 K.

Table A1.2.3.5 CDL description of DT_analysis variable

Storage type	Name	Description	Unit
Byte	DT_analysis	Deviation from previous day (T-1) SSTfnd analysis. If no analysis is available, the reference mean SST climatology should be used defined as R9 in table A3.3.1. Precision : 0.1 K	K
CDL description			
byte DT_analysis (time, nj, ni) ; DT_analysis:long_name = "deviation from sst reference climatology" ; DT_analysis:units = "kelvin" ; DT_analysis:_FillValue = -128b ; DT_analysis:add_offset = 0. ; DT_analysis:scale_factor = 0.1 ; DT_analysis:valid_min = -127b ; DT_analysis:valid_max = 127b ; DT_analysis:coordinates = "lon lat" ; DT_analysis:grid_mapping = "polar_stereographic" ; DT_analysis:reference = "climatology, Faugere and all" ;			
Comments			
Refer to WP-ID2.1.1 and WP-ID2.2.1 for definition The "reference' variable attribute should be used to specify the analysis or climatology used.			

A1.2.3.6 surface_solar_irradiance variable

The variable 'surface_solar_irradiance' will be included with the format requirements shown in Table A1.2.3.6.

Table A1.2.3.6 CDL description of surface_solar_irradiance variable

Storage type	Name	Description	Unit
byte	surface_solar_irradiance	Near contemporaneous integrated SSI value. if 6 hourly irradiance : ranges from 0 to 500 Wm-2, precision 2 Wm-2. if 3 hourly irradiance : ranges from 0 to 1000 Wm-2, precision 4 Wm-2.	Wm-2
CDL description			
byte surface_solar_irradiance(time, nj, ni) ; surface_solar_irradiance:long_name = "surface solar irradiance" ; surface_solar_irradiance:units = "W m-2" ; surface_solar_irradiance:_FillValue = -128b ; surface_solar_irradiance:add_offset = 250. ; surface_solar_irradiance:scale_factor = 0.2 ; surface_solar_irradiance:valid_min = -127b ; surface_solar_irradiance:valid_max = 127b ; surface_solar_irradiance:coordinates = "lon lat" ; surface_solar_irradiance:grid_mapping = "polar_stereographic" ;			
Comments			

Refer to WP-ID2.1.5 for definition The source will be systematically refered in the 'sources_of surface_solar_irradiance' variable instead of using a variable attribute (as stated in the GDS), for consistency reasons across the datasets.

A1.2.3.7 ssi_dtime_from_sst variable

The variable 'ssi_dtime_from_sst' will be included with the format requirements shown in Table A1.2.3.7.

Table A1.2.3.7 CDL description of ssi_dtime_from_sst variable

Storage type	Name	Description	Unit
byte	ssi_dtime_from_sst	time difference of SSI measurement from SST measurement in hours	Scaled hour
CDL description			
byte ssi_dtime_from_sst (time, nj, ni) ; ssi_dtime_from_sst :long_name = "time difference of surface solar irradiance measurement from sst measurement" ; ssi_dtime_from_sst :units = "hour" ; ssi_dtime_from_sst :_FillValue = -128b ; ssi_dtime_from_sst :add_offset = 0. ; ssi_dtime_from_sst :scale_factor = 0.1 ; ssi_dtime_from_sst :valid_min = -127b ; ssi_dtime_from_sst :valid_max = 127b ; ssi_dtime_from_sst :coordinates = "lon lat" ; ssi_dtime_from_sst :grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.5 for definition			

A1.2.3.8 wind_speed variable

The variable 'wind_speed' will be included with the format requirements shown in Table A1.2.3.8.

Table A1.2.3.8 CDL description of wind_speed variable

Storage type	Name	Description	Unit
byte	wind_speed	The wind speed ranges from 0 to 50m/s. precision : 1 m/s.	ms-1
CDL description			
byte wind_speed(time, nj, ni) ; wind_speed:long_name = "wind speed" ; wind_speed:units = "m s-1" ; wind_speed:_FillValue = -128b ; wind_speed:add_offset = 0. ; wind_speed:scale_factor = 1. ; wind_speed:valid_min = -127b ; wind_speed:valid_max = 127b ; wind_speed:coordinates = "lon lat" ; wind_speed:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.4 for definition There is no 'source' attribute or variable anymore			

A1.2.3.9 wspd_dtime_from_sst variable

The variable 'wspd_dtime_from_sst' will be included with the format requirements shown in Table A1.2.3.9.

Table A1.2.3.9 CDL description of wspd_dtime_from_sst variable

Storage type	Name	Description	Unit
byte	wspd_dtime_from_sst	time difference of wind speed measurement from SST measurement in hours. Precision : 0.1 hour	hour
CDL description			
byte wspd_dtime_from_sst me(time, nj, ni) ; wspd_dtime_from_sst :long_name = "time difference of wind speed measurement from sst measurement" ; wspd_dtime_from_sst :units = "hour" ; wspd_dtime_from_sst :_FillValue = -128b ; wspd_dtime_from_sst :add_offset = 0. ; wspd_dtime_from_sst :scale_factor = 0.1 ; wspd_dtime_from_sst :valid_min = -127b ; wspd_dtime_from_sst :valid_max = 127b ; wspd_dtime_from_sst :coordinates = "lon lat" ; wspd_dtime_from_sst :grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.4 for definition			

A1.2.3.10 sea_ice_fraction variable

The variable 'sea_ice_fraction' will be included with the format requirements shown in Table A1.2.3.10. It expresses the fractional sea ice contamination flag. Ranges from 0 to 1, precision 0.01.

Table A1.2.3.10 CDL description of sea_ice_fraction variable

Storage type	Name	Description	Unit
byte	Sea_ice_fraction	fractional sea ice contamination flag. Ranges from 0 to 1, precision 0.01.	%
CDL description			
byte sea_ice_fraction(time, nj, ni) ; sea_ice_fraction:long_name = "sea ice fraction" ; sea_ice_fraction:units = "percent" ; sea_ice_fraction:_FillValue = -128b ; sea_ice_fraction:add_offset = 0. ; sea_ice_fraction:scale_factor = 0.01 ; sea_ice_fraction:valid_min = 0b ; sea_ice_fraction:valid_max = 100b ; sea_ice_fraction:coordinates = "lon lat" ; sea_ice_fraction:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.3 for definition The source will be systematically refered in the 'sources_of_sea_ice_fraction' variable instead of using an variable attribute (as stated in the GDS), for consistency reasons across the datasets			

A1.2.3.11 aerosol_optical_depth variable

The variable 'aerosol_optical_depth' will be included with the format requirements shown in Table A1.2.3.11.

Table A1.2.3.11 CDL description of aerosol_optical_depth variable

Storage type	Name	Description	Unit
byte	aerosol_optical_depth	Aerosol optical depth	none
CDL description			
byte aerosol_optical_depth(time, nj, ni) ; aerosol_optical_depth:long_name = "aerosol optical depth" ; aerosol_optical_depth:units = "count'" ; aerosol_optical_depth:_FillValue = -128b ; aerosol_optical_depth:add_offset = 0. ; aerosol_optical_depth:scale_factor = 1. ; aerosol_optical_depth:valid_min = -127b ; aerosol_optical_depth:valid_max = 127b ; aerosol_optical_depth:coordinates = "lon lat" ; aerosol_optical_depth:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.6 for definition This variable is dimensionless (unit set to 'count' instead of 'micrometer' in GDS). The source will be systematically refered in the 'sources_of_aerosol_optical_depth' variable instead of using an variable attribute (as stated in the GDS), for consistency reasons across the datasets.			

A1.2.3.12 AOD_dtime_from_sst variable

The variable 'AOD_dtime_from_sst' will be included with the format requirements shown in Table A1.2.3.12.

Table A1.2.3.12 CDL description of AOD_dtime_from_sst variable

Storage type	Name	Description	Unit
byte	AOD_dtime_from_sst	time difference of AOD measurement from SST measurement in hours. Precision : 0.1 hour	hour
CDL description			
byte AOD_dtime_from_sst me(time, nj, ni) ; AOD _dtime_from_sst :long_name = "time difference of AOD measurement from sst measurement" ; AOD_dtime_from_sst :units = "hour" ; AOD _dtime_from_sst :_FillValue = -128b ; AOD _dtime_from_sst :add_offset = 0. ; AOD _dtime_from_sst :scale_factor = 0.1 ; AOD _dtime_from_sst :valid_min = -127b ; AOD _dtime_from_sst :valid_max = 127b ; AOD _dtime_from_sst :coordinates = "lon lat" ; AOD _dtime_from_sst :grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.6 for definition			

A1.2.3.13 sources_of_wind_speed variable

The variable 'sources_of_wind_speed' will be included with the format requirements shown in Table A1.2.3.13.

Table A1.2.3.13 CDL description of sources_of_ssi variable

Storage type	Name	Description	Unit
byte	sources_of_wind_speed	Source(s) of wind_speed value	none
CDL description			

byte sources_of_wind_speed (time, nj, ni) ; sources_of_wind_speed:long_name = "sources_of wind_speed" ; sources_of_wind_speed:_FillValue = -128b ; sources_of_wind_speed:comment = "details here source codes" ; sources_of_wind_speed:coordinates = "lon lat" ; sources_of_wind_speed:grid_mapping = "polar_stereographic" ;
Comments
Refer to WP-ID2.1.4 for definition NetCDF minimum size for a variable is the byte. The GDS requirement for 4 bits to encode the source can not be satisfied with netCDF : the source will thus be encoded on a full byte. The source codes must be detailed in the 'comment' attribute

the source code is selected from Table 2.1.4

A1.2.3.14 sources_of_ssi variable

The variable 'sources_of_ssi' will be included with the format requirements shown in Table A1.2.3.14.

Table A1.2.3.14 CDL description of sources_of_ssi variable

Storage type	Name	Description	Unit
byte	sources_of_ssi	Source(s) of SSI value	none
CDL description			
byte sources_of_ssi(time, nj, ni) ; sources_of_ssi:long_name = "sources_of surface solar irradiance" ; sources_of_ssi:_FillValue = -128b ; sources_of_ssi:comment = "details here source codes" ; sources_of_ssi:coordinates = "lon lat" ; sources_of_ssi:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.5 for definition NetCDF minimum size for a variable is the byte. The GDS requirement for 4 bits to encode the source can not be satisfied with netCDF : the source will thus be encoded on a full byte. The source codes must be detailed in the 'comment' attribute			

the source code is selected from Table 2.1.5

A1.2.3.15 sources_of sea_ice_fraction variable

The variable 'sources_of sea_ice_fraction' will be included with the format requirements shown in Table A1.2.3.15.

Table A1.2.3.15 CDL description of sources_of sea_ice_fraction variable

Storage type	Name	Description	Unit
byte	sources_of_sea_ice_fraction	Source(s) of sea ice values	none
CDL description			
byte sources_of sea_ice_fraction (time, nj, ni) ; sources_of sea_ice_fraction:long_name = "sources of sea ice fraction " ; sources_of sea_ice_fraction:_FillValue = -128b ; sources_of sea_ice_fraction:comment = "details here source codes" ; sources_of sea_ice_fraction:coordinates = "lon lat" ; sources_of sea_ice_fraction:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.3 for definition NetCDF minimum size for a variable is the byte. The GDS requirement for 4 bits to encode the source can not be satisfied with netCDF : the source will thus be encoded on a full byte			

The source codes must be detailed in the 'comment' attribute

the source code is selected from Table 2.1.3.

A1.2.3.16 sources_of aerosol_optical_depth variable

The variable 'sources_of aerosol_optical_depth' will be included with the format requirements shown in Table A1.2.3.16.

Table A1.2.3.16 CDL description of sources_of aerosol_optical_depth variable

Storage type	Name	Description	Unit
byte	sources_of_AOD	Source(s) of AOD values	none
CDL description			
byte sources_of_AOD (time, nj, ni) ; sources_of_AOD:long_name = "sources of aerosol optical depth " ; sources_of_AOD:_FillValue = -128b ; sources_of_AOD:comment = "details here source codes" ; sources_of_AOD:coordinates = "lon lat" ; sources_of_AOD:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.6 for definition NetCDF minimum size for a variable is the byte. The GDS requirement for 4 bits to encode the source can not be satisfied with netCDF : the source will thus be encoded on a full byte The source codes must be detailed in the 'comment' attribute			

the source code is selected from Table 2.1.6.

A1.2.3.17 satellite_zenith_angle variable

The variable 'satellite_zenith_angle' will be included with the format requirements shown in Table A1.2.3.17.

Table A1.2.3.17 CDL description of satellite_zenith_angle variable

Storage type	Name	Description	Unit
byte	satellite_zenith_angle	Ranges from -90 to 90. Precision : 1 degree.	degree
CDL description			
byte satellite_zenith_angle(time, nj, ni) ; satellite_zenith_angle:long_name = "satellite zenith angle" ; satellite_zenith_angle:units = "angular_degree" ; satellite_zenith_angle:_FillValue = -128b ; satellite_zenith_angle:add_offset = 0. ; satellite_zenith_angle:scale_factor = 1. ; satellite_zenith_angle:valid_min = -90b ; satellite_zenith_angle:valid_max = 90b ; satellite_zenith_angle:coordinates = "lon lat" ; satellite_zenith_angle:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.1.7 for definition			

A1.2.3.18 rejection_flag variable

The variable 'rejection_flag' will be included with the format requirements shown in Table A1.2.3.18.

Table A1.2.3.18 CDL description of rejection_flag variable

Storage	Name	Description	Unit

type			
byte	rejection_flag	Rejection flag	none
CDL description			

byte rejection_flag(time, nj, ni) ;
rejection_flag:long_name = "rejection flag" ;
rejection_flag:comment = "
b0 : 1= SST out of range;
b1 : 1= Cosmetic value;
b2 : 1= IR cloudy
b3 : 1= MW rain
b4 : 1= ice
b5 : 1= spare
b6 : 1= Land
b7 : 1= spare
rejection_flag:coordinates = "lon lat" ;
rejection_flag:grid_mapping = "polar_stereographic" ;			

Comments			

The meaning of each bit must be detailed in the 'comment' variable attribute.
b0 : 1= SST out of range;
b1 : 1= Cosmetic value;
b2 : 1= IR cloudy
b3 : 1= MW rain
b4 : 1= MW ice
b5 : 1= MW wind
b6 : 1= Land;
b7 : 1= spare			

A1.2.3.19 confidence_flag variable

The variable 'confidence_flag' will be included with the format requirements shown in Table A1.2.3.19.

Table A1.2.3.19 CDL description of confidence_flag variable

Storage type	Name	Description	Unit
byte	confidence_flag	Confidence flag	none
CDL description			
byte confidence_flag(time, nj, ni) ; confidence_flag:long_name = "confidence flag" ; confidence_flag:comment = " b0 : 1=potential side lobe contamination; b1 : 1=relaxed rain contamination suspected; b2 : 1=TMI SST retrieved in SST < 285K b3 : 1=high wind speed retrieval b4 : 1=sea ice retrieval for MW data b5 : 1= sun glint suspected b6 : 1= L2 native bias and standard deviation; b7 : 1= L2 native confidence value " ; confidence_flag:coordinates = "lon lat" ; confidence_flag:grid_mapping = "polar_stereographic" ;			
Comments			
The meaning of each bit must be detailed in the 'comment' variable attribute. This comes from the GDS table 1.2.3 which is not consistent with the CDL description of a L2P data file. Some confidence flag fields were assigned to specific variables (like ancillary data sources and delta times, proximity confidence). The L2P_version field was assigned to the global			

attribute ‘product_version’. b0 : 1=potential side lobe contamination; b1 : 1=relaxed rain contamination suspected; b2 : 1=TMI SST retrieved in SST < 285K b3 : 1=high wind speed retrieval b4 : 1=sea ice retrieval for MW data b5 : 1= sun glint suspected b6 : 1= L2 native bias and standard deviation; b7 : 1= L2 native confidence value " ;

A1.2.3.20 proximity_confidence variable

The variable ‘proximity_confidence’ will be included with the format requirements shown in Table A1.2.3.20.

Table A1.2.3.20 CDL description of proximity_confidence variable

Storage type	Name	Description	Unit
byte	proximity_confidence’	proximity confidence value	none
CDL description			
byte proximity_confidence’ (time, nj, ni) ; proximity_confidence’:long_name = "proximity confidence value" ; proximity_confidence’:_FillValue = -128b ; proximity_confidence’:coordinates = "lon lat" ; proximity_confidence’:grid_mapping = "polar_stereographic" ;			
Comments			
Refer to WP-ID2.2.1.5 and WP-ID2.2.2.7 for definition			

A1.2.3 Sample L2P file (CDL header)

A complete CDL description of a L2P data file is given in Figure A1.2.3.1.

```
netcdf example {
dimensions:
        ni = 1024 ;
        nj = 1024 ;
        time = 1 ;

variables:
        float lat(nj, ni) ;
                lat:long_name = "latitude" ;
                lat:units = "degrees_north" ;
                lat:add_offset = 20. ;
                lat:scale_factor = 0.001 ;
        float lon(nj, ni) ;
                lon:long_name = "longitude" ;
                lon:units = "degrees_east" ;
                lon:add_offset = 0. ;
                lon:scale_factor = 0.001 ;
        long time(time) ;
                time:long_name = "reference time of sst file" ;
                time:units = "seconds since 1981-01-01 00:00:00" ;
        char polar_stereographic ;
                polar_stereographic:grid_mapping_name = "polar_stereographic" ;
                polar_stereographic:straight_vertical_longitude_from_pole = 0. ;
                polar_stereographic:scale_factor_at_projection_origin = 1.0 ;
                polar_stereographic:false_easting = 0. ;
                polar_stereographic:false_northing = 0. ;
        short sea_surface_temperature(time, nj, ni) ;
                sea_surface_temperature:long_name = "sea surface temperature" ;
                sea_surface_temperature:units = "kelvin" ;
                sea_surface_temperature:_FillValue = -32768s ;
                sea_surface_temperature:add_offset = 273.15 ;
                sea_surface_temperature:scale_factor = 0.01 ;
                sea_surface_temperature:valid_min = -5000s ;
```

```
                sea_surface_temperature:valid_max = 5000s ;
                sea_surface_temperature:coordinates = "lon lat" ;
                sea_surface_temperature:grid_mapping = "polar_stereographic" ;
                sea_surface_temperature:source = " name of L2 source  (ex:EUMETSAT SAF O&SI)" ;
        short sst_dtime (time, nj, ni) ;
                sst_dtime:long_name = "time difference from reference time" ;
                sst_dtime:units = "second" ;
                sst_dtime:_FillValue = -32768s ;
                sst_dtime:add_offset = 0. ;
                sst_dtime:scale_factor = 1. ;
                sst_dtime:valid_min = -32767s ;
                sst_dtime:valid_max = 32767s ;
                sst_dtime:coordinates = "lon lat" ;
                sst_dtime:grid_mapping = "polar_stereographic" ;
        byte SSES_bias_error (time, nj, ni) ;
                SSES_bias_error:long_name = "SSES bias error based on confidence flags" ;
                SSES_bias_error:units = "kelvin" ;
                SSES_bias_error:_FillValue = -128 ;
                SSES_bias_error:add_offset = 0. ;
                SSES_bias_error:scale_factor = 0.01 ;
                SSES_bias_error:valid_min = -127 ;
                SSES_bias_error:valid_max = 127 ;
                SSES_bias_error:coordinates = "lon lat" ;
                SSES_bias_error:grid_mapping = "polar_stereographic" ;
        byte SSES_standard_deviation_error (time, nj, ni) ;
                SSES_standard_deviation_error:long_name = "SSES standard deviation error based on confidence
flags" ;
                SSES_standard_deviation_error:units = "kelvin" ;
                SSES_standard_deviation_error:_FillValue = -128 ;
                SSES_standard_deviation_error:add_offset = 100. ;
                SSES_standard_deviation_error:scale_factor = 0.01 ;
                SSES_standard_deviation_error:valid_min = -127 ;
                SSES_standard_deviation_error:valid_max = 127 ;
                SSES_standard_deviation_error:coordinates = "lon lat" ;
                SSES_standard_deviation_error:grid_mapping = "polar_stereographic" ;
        byte DT_analysis (time, nj, ni) ;
                DT_analysis:long_name = "deviation from sst reference climatology" ;
                DT_analysis:units = "kelvin" ;
                DT_analysis:_FillValue = -128 ;
                DT_analysis:add_offset = 0. ;
                DT_analysis:scale_factor = 0.1 ;
                DT_analysis:valid_min = -127 ;
                DT_analysis:valid_max = 127 ;
                DT_analysis:coordinates = "lon lat" ;
                DT_analysis:grid_mapping = "polar_stereographic" ;
                DT_analysis:reference = "climatology, Faugere and all" ;
        byte surface_solar_irradiance(time, nj, ni) ;
                surface_solar_irradiance:long_name = "surface solar irradiance" ;
                surface_solar_irradiance:units = "W m-2" ;
                surface_solar_irradiance:_FillValue = -128 ;
                surface_solar_irradiance:add_offset = 250. ;
                surface_solar_irradiance:scale_factor = 0.2 ;
                surface_solar_irradiance:valid_min = -127 ;
                surface_solar_irradiance:valid_max = 127 ;
                surface_solar_irradiance:coordinates = "lon lat" ;
                surface_solar_irradiance:grid_mapping = "polar_stereographic" ;
        byte ssi_dtime_from_sst(time, nj, ni) ;
                ssi_dtime_from_sst:long_name = "time difference of surface solar irradiance measurement from sst
measurement" ;
                ssi_dtime_from_sst:units = "hour" ;
                ssi_dtime_from_sst:_FillValue = -128 ;
                ssi_dtime_from_sst:add_offset = 0. ;
                ssi_dtime_from_sst:scale_factor = 0.1 ;
                ssi_dtime_from_sst:valid_min = -127 ;
                ssi_dtime_from_sst:valid_max = 127 ;
                ssi_dtime_from_sst:coordinates = "lon lat" ;
                ssi_dtime_from_sst:grid_mapping = "polar_stereographic" ;
        byte wind_speed(time, nj, ni) ;
                wind_speed:long_name = "wind speed" ;
                wind_speed:units = "m s-1" ;
                wind_speed:_FillValue = -128 ;
                wind_speed:add_offset = 0. ;
                wind_speed:scale_factor = 1. ;
                wind_speed:valid_min = -127 ;
                wind_speed:valid_max = 127 ;
```

```
                wind_speed:coordinates = "lon lat" ;
                wind_speed:grid_mapping = "polar_stereographic" ;
        byte wspd_dtime_from_sst(time, nj, ni) ;
                wspd_dtime_from_sst:long_name = "time difference of wind speed measurement from sst
measurement" ;
                wspd_dtime_from_sst:units = "hour" ;
                wspd_dtime_from_sst:_FillValue = -128 ;
                wspd_dtime_from_sst:add_offset = 0. ;
                wspd_dtime_from_sst:scale_factor = 0.1 ;
                wspd_dtime_from_sst:valid_min = -127 ;
                wspd_dtime_from_sst:valid_max = 127 ;
                wspd_dtime_from_sst:coordinates = "lon lat" ;
                wspd_dtime_from_sst:grid_mapping = "polar_stereographic" ;
        byte sea_ice_fraction(time, nj, ni) ;
                sea_ice_fraction:long_name = "sea ice fraction" ;
                sea_ice_fraction:units = "percent" ;
                sea_ice_fraction:_FillValue = -128 ;
                sea_ice_fraction:add_offset = 0. ;
                sea_ice_fraction:scale_factor = 0.01 ;
                sea_ice_fraction:valid_min = 0 ;
                sea_ice_fraction:valid_max = 100 ;
                sea_ice_fraction:coordinates = "lon lat" ;
                sea_ice_fraction:grid_mapping = "polar_stereographic" ;
        byte aerosol_optical_depth(time, nj, ni) ;
                aerosol_optical_depth:long_name = "aerosol optical depth" ;
                aerosol_optical_depth:units = "count'" ;
                aerosol_optical_depth:_FillValue = -128 ;
                aerosol_optical_depth:add_offset = 0. ;
                aerosol_optical_depth:scale_factor = 1. ;
                aerosol_optical_depth:valid_min = -127 ;
                aerosol_optical_depth:valid_max = 127 ;
                aerosol_optical_depth:coordinates = "lon lat" ;
                aerosol_optical_depth:grid_mapping = "polar_stereographic" ;
        byte sources_of_ssi(time, nj, ni) ;
                sources_of_ssi:long_name = "sources_of surface solar irradiance" ;
                sources_of_ssi:_FillValue = -128 ;
                sources_of_ssi:comment = "details here source codes" ;
                sources_of_ssi:coordinates = "lon lat" ;
                sources_of_ssi:grid_mapping = "polar_stereographic" ;
        byte sources_of_sea_ice_fraction (time, nj, ni) ;
                sources_of_sea_ice_fraction:long_name = "sources of sea ice fraction " ;
                sources_of_sea_ice_fraction:_FillValue = -128 ;
                sources_of_sea_ice_fraction:comment = "details here source codes" ;
                sources_of_sea_ice_fraction:coordinates = "lon lat" ;
                sources_of_sea_ice_fraction:grid_mapping = "polar_stereographic" ;
        byte sources_of_aod(time, nj, ni) ;
                sources_of_aod:long_name = "sources of aerosol optical depth " ;
                sources_of_aod:_FillValue = -128 ;
                sources_of_aod:comment = "details here source codes" ;
                sources_of_aod:coordinates = "lon lat" ;
                sources_of_aod:grid_mapping = "polar_stereographic" ;
        byte satellite_zenith_angle(time, nj, ni) ;
                satellite_zenith_angle:long_name = "satellite zenith angle" ;
                satellite_zenith_angle:units = "angular_degree" ;
                satellite_zenith_angle:_FillValue = -128 ;
                satellite_zenith_angle:add_offset = 0. ;
                satellite_zenith_angle:scale_factor = 1. ;
                satellite_zenith_angle:valid_min = -90 ;
                satellite_zenith_angle:valid_max = 90 ;
                satellite_zenith_angle:coordinates = "lon lat" ;
                satellite_zenith_angle:grid_mapping = "polar_stereographic" ;
        byte rejection_flag(time, nj, ni) ;
                rejection_flag:long_name = "rejection flag" ;
                rejection_flag:comment = "
b0 : 1= SST out of range;
b1 : 1= Cosmetic value;
b2 : 1= IR cloudy
b3 : 1= MW rain
b4 : 1= ice
b5 : 1= spare
b6 : 1= Land;
b7 : 1= spare " ;
                rejection_flag:coordinates = "lon lat" ;
                rejection_flag:grid_mapping = "polar_stereographic" ;
        byte confidence_flag(time, nj, ni) ;
```

```
                              confidence_flag:long_name = "confidence flag" ;
                              confidence_flag:comment = "
b0 : 1=potential side lobe contamination;
b1 : 1=relaxed rain contamination suspected;
b2 : 1=TMI SST retrieved in SST < 285K
b3 : 1=high wind speed retrieval
b4 : 1=sea ice retrieval for MW data
b5 : 1= sun glint suspected
b6 : 1= L2 native bias and standard deviation;
b7 : 1= L2 native confidence value " ;
                              confidence_flag:coordinates = "lon lat" ;
                              confidence_flag:grid_mapping = "polar_stereographic" ;
              byte proximity_confidence_value(time, nj, ni) ;
                              proximity_confidence_value:long_name = "proximity confidence value" ;
                              proximity_confidence_value:_FillValue = -128 ;
                              proximity_confidence_value:coordinates = "lon lat" ;
                              proximity_confidence_value:grid_mapping = "polar_stereographic" ;

// global attributes:
                              :Conventions = "CF-1.0" ;
                              :title = "Sea Surface Temperature from AVHRR onboard NOAA-16, 2km resolution, over West
Mediterranean sea" ;
                              :DSD_entry_id = "EUR-CMS-AVHRR16_L-IR-L2P-MOCC" ;
                              :references = "Medspiration products user manual, Robinson I., Leborgne P., Piolle J.F., Larnicol G.,
v1.02, September 2004" ;
                              :institution = "MEDSPIRATION" ;
                              :contact = "Pierre.Leborgne@meteo.fr" ;
                              :GDS_version_id = "v1.0-rev1.4" ;
                              :netcdf_version_id = "3.5" ;
                              :creation_date = "2004-08-25" ;
                              :product_version = "1.0" ;
                              :history = " " ;
                              :platform = "NOAA-16" ;
                              :sensor = "avhrr" ;
                              :spatial_resolution = "2 km" ;
                              :start_date = "2004-08-25 UTC" ;
                              :start_time = "00:12:23 UTC" ;
                              :stop_date = "2004-08-25 UTC" ;
                              :stop_time = "00:14:18 UTC" ;
                              :southernmost_latitude = "29.08f" ;
                              :northernmost_latitude = "47.51f" ;
                              :westernmost_longitude = "-8.32f" ;
                              :easternmost_longitude = "18.85f" ;
                              :file_quality_index = "0" ;
                              :comment = " " ;
}
```

Figure A1.2.3.1 CDL description of an example L2P data file.

A1.3. Analysed L4 (L4) SST data products

L4 analysed data products, L4 are derived from an analysis procedure based on L2P data products. L4 data products are produced for every APPW. L4 data products are formatted as netCDF format data files and include ancillary data and error statistics for each grid-cell. The UHR/L4 format requirements for the global/variable attributes and the coordinate variables are equivalent to L2P products format requirements (netCDF format, complying to CF-1.0 convention).

A1.3.1 L4 product filename convention

The GDS filename convention used for L4 data products has been designed to provide useful information in an easily readable format. All L4 data product filenames are derived according to the following convention:

> **<Date Valid>-<Processing Centre Code>-L4<Product type>-<Area>-<Processing Model ID>.<base format>**

which is defined in Table A1.3.1.

Table A1.3.1. L4 analysed data product filename components.

Name	Definition	Description
<Processing Centre Code>	Refer to Appendix A2 Table A2.1	Processing centre code
<Area>	Table A1.3.2	The area covered by the L4 product
<Date Valid>	YYYYMMDD	Refers to the date for which this particular data set
<product type>	LRfnd=low resolution, UHfnd=ultra-high resolution	Resolution of analysed foundation SST (fnd) data
<processing model ID>	v*nn* (where *nn* is the GDS version number, e.g., 01	Version number of the GDS system used to process the data file
<base format>	Nc	Generic file format (nc=netCDF)

For example:

`20040621-EUR-L4UHfnd-MED-v01.nc`

Refers to a data set valid on 21st June 2004 (20040621) generated at the European RDAC (EUR), the data is an estimate of the foundation SST at ultra-high resolution(L4UHfnd) covering the Mediterranean area (MED), it was generated using GDS version 1 (v01) and is formatted as a GHRSST-PP netCDF file (.nc)

Table A1.3.2. L4 data product filename area code definitions. (Rev 1, 26/02/2004)

Code	Definition	Description
GLOB	90°S to 90°N and from 180°W - 180°E	Global coverer age data sets
MED	30°N to 46°N and from 6°W to 36.5°E	Mediterranean sea area
EURDAC	70°S to 90°N and from 100°W to 45°E	European RDAC area served by the ESA Medspiration project
NORSEA	48°N to 75°N and from 12°W to 70°E	Nordic Seas area
BLKSEA	40°N to 48°N and from 27°E to 42°E	Black Sea area
…		

A1.3.2. L4SST and UHRSSTfnd grid cell ancillary data record format specification

Table A1.3.2 describes format of GDS L4 grid cell ancillary data that should be created for each L4 grid cell.

Table A1.3.2 GDSv1 generic L4 SST product data product (v4.0 12/01/04) agreed at the 4th GHRSST-PP Science Team Meeting., Los Angeles, USA, September 2003 and during subsequent discussions. Mandatory fields are shown with blue shading.

Name	Definition in GDS-v1.0	Description	Units
SST_foundation	WP-ID5	SSTfnd from analysis system e	K, scaled int
normalised_analysis_error	WP-ID3.2.3	Error estimate output from the analysis system	K, scaled int
Bias	Rule 6.2.3	Analysis error bias	K, Scaled INT
Sea_ice_fraction	WP-ID2.1.1.3	Fractional Sea Ice	Percent

		concentration	coverage
DT_SSTskin1:T=00:00UTC **DT_SSTskin2:T=02:00UTC** **DT_SSTskin3:T=04:00UTC** **DT_SSTskin4:T=06:00UTC** **DT_SSTskin5:T=08:00UTC** **DT_SSTskin6:T=10:00UTC** **DT_SSTskin7:T=12:00UTC** **DT_SSTskin8:T=14:00UTC** **DT_SSTskin9:T=16:00UTC** **DT_SSTskin10:T=18:00UTC** **DT_SSTskin11:T=20:00UTC** **DT_SSTskin12:T=22:00UTC**	WP-ID4.4	Best estimate of the mean SSTskin over a 2 hour period expressed as a deviation from SSTfnd at 12 synoptic times 1 value per byte.	K, scaled
skin_parameterisation_scheme	WP-ID4.3	Code indicating the skin parameterisation scheme used to derive 12 SSTskin estimates. In the GDS v1.0 this may be defined by the RDAC/GDAC analysis system producing the SSTfnd data sets.	Code
sources_of sea_ice_fraction	WP-ID2.1.1.3	Source of Fractional sea ice data should be same definition as for L2P	Code
SST_skin1_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin2_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin3_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin4_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin5_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin6_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain	Code

		3=Poor	
SST_skin7_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin8_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin9_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin10_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin11_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
SST_skin12_QC	WP-ID4.4	Quality control indicator: 0=Good 1=Fair 2=Uncertain 3=Poor	Code
mask	WP-ID2.1.1.11	0=sea, 1=land, 2=lakes, ice=3	Flag

A1.3.3 SSTfnd/UHRSSTfnd L2P product file Common data format Description Language (CDL)

The following sections provide a detailed reference for GHRSST-PP SSTfnd/UHRSSTfnd netCDF data files.

A1.3.3.1 Coordinate variables

A 'skin-time' coordinate variable (and dimension) is required to provide a time axis for the 'skin_sea_surface_temperature' variable. It will contain 12 dates, every 2 hours starting from 00h00 to 22h00.The dimensions and coordinate variables shall be formatted as follows:

```
dimensions:
  lon = 1024 ;
  lat = 1024 ;
  time = 1 ;
  skin_time = 12 ;

variables:
 long time(time) ;
```

```
    time:long_name = "reference time of sst field" ;
    time:units = "seconds since 1981-01-01 00:00:00" ;
 long skin_time(time) ;
    time:long_name = "time of skin sst field" ;
    time:units = "seconds since 1981-01-01 00:00:00" ;
 float lat(lat) ;
    lat:long_name = "latitude" ;
    lat:units = "degrees_north" ;
 float lon(lon) ;
    lon:long_name = "longitude" ;
    lon:units = "degrees_east" ;
```

A1.3.3.2 sst_foundation variable

The variable 'sst_foundation' will be included with the format requirements shown in Table A1.3.3.2.

Table A1.3.3.2 CDL description of sst_foundation variable

Storage type	Name	Description	Unit
short	sst_foundation	SSTfnd from analysis	K
CDL description			
short sst_foundation(time, lat, lon) ; sst_foundation:long_name = "foundation sea surface temperature" ; sst_foundation:units = "kelvin" ; sst_foundation:_FillValue = -32768s ; sst_foundation:add_offset = 273.15 ; sst_foundation:scale_factor = 0.01 ; sst_foundation:valid_min = -32767s ; sst_foundation:valid_max = 32767s ;			
Comments			
Refer to WP-ID5 for definition			

A1.3.3.3 normalised_analysis_error variable

The variable 'normalised_analysis_error' will be included with the format requirements shown in Table A1.3.3.3.

Table A1.3.3.3 CDL description of normalised_analysis_error variable

Storage type	Name	Description	Unit
byte	normalised_analysis_error	Error estimate output from analysis system	K
CDL description			

byte normalised_analysis_error(time, lat, lon) ;
normalised_analysis_error:long_name = "error estimate from the analysis" ;
normalised_analysis_error:units = "kelvin" ;
normalised_analysis_error:_FillValue = -128;
normalised_analysis_error:add_offset = 0. ;
normalised_analysis_error:scale_factor = 0.01 ;
normalised_analysis_error:valid_min = -127;
normalised_analysis_error:valid_max = 127;			

Comments			
Refer to WP-ID3.2.3 for definition			

A1.3.3.4 bias variable

The variable 'bias' will be included with the format requirements shown in Table A1.3.3.4.

Table A1.3.3.4 CDL description of bias variable

Storage type	Name	Description	Unit
byte	bias	Analysis error bias	K
CDL description			
byte bias(time, lat, lon) ; bias:long_name = "analysis error bias" ; bias:units = "kelvin" ; bias:_FillValue = -128 ; bias:add_offset = 0. ; bias:scale_factor = 0.01 ; bias:valid_min = -127 ; bias:valid_max = 127 ;			
Comments			
Refer to rule 6.2.3 for definition			

A1.3.3.5 sea_ice_fraction variable

The variable 'sea_ice_fraction' will be included with the format requirements shown in Table A1.3.3.5.

Table A1.3.3.5 CDL description of sea_ice_fraction variable

Storage type	Name	Description	Unit
byte	sea_ice_fraction	Fractional sea ice concentration	%
CDL description			
byte sea_ice_fraction(time, lat, lon) ; sea_ice_fraction:long_name = "sea ice fraction" ; sea_ice_fraction:units = "percent" ; sea_ice_fraction:_FillValue = -128 ; sea_ice_fraction:add_offset = 0. ; sea_ice_fraction:scale_factor = 0.01 ; sea_ice_fraction:valid_min = -127 ; sea_ice_fraction:valid_max = 127 ;			
Comments			
Refer to WP-ID2.1.1.3 for definition This variable shall not be used for the Mediterranean UHR/L4 product or any other area for which this parameter is not relevant. The 'source' variable attribute was removed : the variable 'sources_of sea_ice_fraction' was added instead to reflect the GDS table A1.3.2 specifications.			

A1.3.3.6 sources_of sea_ice_fraction variable

The variable 'sources_of sea_ice_fraction' will be included if relevant with the format requirements shown in Table A1.3.3.6.

Table A1.3.3.6 CDL description of sources_of sea_ice_fraction variable

Storage type	Name	Description	Unit
byte	sources_of sea_ice_fraction	Source(s) of fractional sea ice concentration (should be the same as for L2P)	none
CDL description			
byte sources_of_sea_ice_fraction (time, lat, lon) ; sources_of_sea_ice_fraction:long_name = "sources of sea ice fraction " ; sources_of_sea_ice_fraction:_FillValue = -128b ;			

sources_of_sea_ice_fraction:comment = "details here source codes" ;
Comments
Refer to WP-ID2.1.1.3 for definition This variable shall not be used for the Mediterranean UHR/L4 product or any other area for which this parameter is not relevant.

the source code is selected from Table 2.1.3.

A1.3.3.7 mask variable

The variable 'mask' will be included with the format requirements shown in Table A1.3.3.7.

Table A1.3.3.7 CDL description of mask variable

Storage type	Name	Description	Unit
byte	mask	Source(s) of fractional sea ice concentration (should be the same as for L2P)	none
CDL description			
byte mask(time, lat, lon) ; mask:long_name = "mask " ; mask:_FillValue = -128b ; mask:comment = " b0: 1=sea b1: 1=land b2: 1=lakes b3: 1=ice;			
Comments			
Refer to WP-ID2.1.1.11 for definition b0: 1=sea b1: 1=land b2: 1=lakes b3: 1=ice			

A1.3.3.8 DT_sst_skin variable

The variable 'DT_sst_skin' will be included with the format requirements shown in Table A1.3.3.8.

Table A1.3.3.8 CDL description of DT_sst_skin variable

Storage type	Name	Description	Unit
byte	DT_sst_skin	Best estimate of the SST skin over a 2-hour period expressed as a deviation from SST foundation at 12 synoptic times. Precision : 0.1 K	K
CDL description			
byte DT_sst_skin (skin_time, lat, lon) ; DT_sst_skin:long_name = "skin sea surface temperature" ; DT_sst_skin:units = "kelvin" ; DT_sst_skin:_FillValue = -128; DT_sst_skin:add_offset = 0; DT_sst_skin:scale_factor = 0.1 ; DT_sst_skin:valid_min = -127; DT_sst_skin:valid_max = 127; DT_sst_skin:source="Free text based on table A1.3.2.";			
Comments			

Refer to WP-ID4.4 for definition The 'source' variable attribute was introduced instead of using a specific variable 'Skin_parameterisation_scheme' as suggested in the GDS table A.1.3.2.

A1.3.3.9 sst_skin_quality_flag variable

The variable 'sst_skin_quality_flag' will be included with the format requirements shown in Table A1.3.3.9.

Table A1.3.3.9 CDL description of sst_skin_quality_flag variable

Storage type	Name	Description	Unit
long	sst_skin_quality_flag	quality control indicator for sst skin	none
CDL description			
byte sst_skin_quality_flag (lat, lon) ; sst_skin_quality_flag:long_name = "quality control indicator for sst skin" ; sst_skin_quality_flag:_FillValue = -128b ; sst_skin_quality_flag:comment = " each pair of bits is related to one of the 12 DT_sst_skin values, in chronological order, with the following code convention : bit(i+1,i): 00 = good, 01 = fair, 10 = uncertain, 11 = poor";			
Comments			
Refer to WP-ID4.4 for definition The SSTskin quality flag consists in pairs of bits, each pair related to one of the 12 DT_sst_skin values (in chronological order). For instance, bits 0 and 1 are related to the DT_sst_skin value at 00:00. For each pair, the following codes are used : (b_{i+1}, b_i) : 00 = good (b_{i+1}, b_i) : 01 = fair (b_{i+1}, b_i) : 10 = uncertain (b_{i+1}, b_i) : 11 = poor			

A1.3.4 Sample UHR/L4 file (CDL header)

A complete CDL description of a UHRSSTfnd data file is given in Figure A1.3.4.1.

```
netcdf example {
dimensions:
        lon = 1024 ;
        lat = 1024 ;
        time = 1 ;
        skin_time = 12 ;

variables:
        long time(time) ;
                time:long_name = "reference time of sst field" ;
                time:units = "seconds since 1981-01-01 00:00:00" ;
        long skin_time(time) ;
                skin_time:long_name = "time of skin sst field" ;
                skin_time:units = "seconds since 1981-01-01 00:00:00" ;
        float lat(lat) ;
                lat:long_name = "latitude" ;
                lat:units = "degrees_north" ;
        float lon(lon) ;
                lon:long_name = "longitude" ;
                lon:units = "degrees_east" ;
        short sst_foundation(time, lat, lon) ;
                sst_foundation:long_name = "foundation sea surface temperature" ;
                sst_foundation:units = "kelvin" ;
                sst_foundation:_FillValue = -32768s ;
                sst_foundation:add_offset = 273.15 ;
                sst_foundation:scale_factor = 0.01 ;
                sst_foundation:valid_min = -32767s ;
```

```
                    sst_foundation:valid_max = 32767s ;
            byte normalised_analysis_error(time, lat, lon) ;
                    normalised_analysis_error:long_name = "error estimate from the analysis" ;
                    normalised_analysis_error:units = "kelvin" ;
                    normalised_analysis_error:_FillValue = -128;
                    normalised_analysis_error:add_offset = 0. ;
                    normalised_analysis_error:scale_factor = 0.01 ;
                    normalised_analysis_error:valid_min = -127;
                    normalised_analysis_error:valid_max = 127;
            byte bias(time, lat, lon) ;
                    bias:long_name = "analysis error bias" ;
                    bias:units = "kelvin" ;
                    bias:_FillValue = -128 ;
                    bias:add_offset = 0. ;
                    bias:scale_factor = 0.01 ;
                    bias:valid_min = -127 ;
                    bias:valid_max = 127 ;
            byte sea_ice_fraction(time, lat, lon) ;
                    sea_ice_fraction:long_name = "sea ice fraction" ;
                    sea_ice_fraction:units = "percent" ;
                    sea_ice_fraction:_FillValue = -128 ;
                    sea_ice_fraction:add_offset = 0. ;
                    sea_ice_fraction:scale_factor = 0.01 ;
                    sea_ice_fraction:valid_min = -127 ;
                     sea_ice_fraction:valid_max = 127 ;
            byte sources_of_sea_ice_fraction (time, lat, lon) ;
                    sources_of_sea_ice_fraction:long_name = "sources of sea ice fraction " ;
                    sources_of_sea_ice_fraction:_FillValue = -128 ;
                    sources_of_sea_ice_fraction:comment = "details here source codes" ;
            byte mask(time, lat, lon) ;
                    mask:long_name = "mask " ;
                    mask:_FillValue = -128 ;
                    mask:comment = "
b0: 1=sea
b1: 1=land
b2: 1=lakes
b3: 1=ice" ;
            byte DT_sst_skin (skin_time, lat, lon) ;
                    DT_sst_skin:long_name = "skin sea surface temperature" ;
                    DT_sst_skin:units = "kelvin" ;
                    DT_sst_skin:_FillValue = -128;
                    DT_sst_skin:add_offset = 0;
                    DT_sst_skin:scale_factor = 0.1 ;
                    DT_sst_skin:valid_min = -127;
                    DT_sst_skin:valid_max = 127;
                    DT_sst_skin:source="Stuart-Menteth model";
            long sst_skin_quality_flag (lat, lon) ;
                    sst_skin_quality_flag:long_name = "quality control indicator for sst skin" ;
                    sst_skin_quality_flag:_FillValue = -2147483648 ;
                    sst_skin_quality_flag:comment = "each pair of bits is related to one of the 12 DT_sst_skin values, in
chronological order, with the following code convention : bit(i+1,i): 00 = good, 01 = fair, 10 = uncertain, 11 = poor" ;

// global attributes:
                    :Conventions = "CF-1.0" ;
                    :title = "Analyzed foundation sea surface temperature over Mediterranean sea" ;
                    :DSD_entry_id = "EUR-L4UHfnd-v01-MED" ;
                    :references = "Medspiration products user manual, Robinson I., Leborgne P., Piolle J.F., Larnicol G.,
v1.02, September 2004" ;
                    :institution = "MEDSPIRATION" ;
                    :contact = "Pierre.Leborgne@meteo.fr" ;
                    :GDS_version_id = "v1.0-rev0.99" ;
                    :netcdf_version_id = "3.5" ;
                    :creation_date = "2004-08-25" ;
                    :product_version = "1.0" ;
                    :history = " " ;
                    :grid_resolution = "0.02 degree" ;
                    :start_date = "2004-08-25" ;
                    :start_time = "00:00:00" ;
                    :stop_date = "2004-08-26" ;
                    :stop_time = "00:00:00" ;
                    :southernmost_latitude = "30.00f" ;
                    :northernmost_latitude = "46.00f" ;
                    :westernmost_longitude = "-8.00f" ;
                    :easternmost_longitude = "40.00f" ;
                    :file_quality_index = "0" ;
```

```
                    :comment = " " ;

        }
```

Figure 1.3.4.1 CDL for a GHRSST-PP L4SSTfnd AVHRR/SST data PRODUCT over the Mediterranean sea product (26/02/04).

Appendix A2. GDS reference code tables

The following sections provide reference tables that are used within the GDS.

A2.1 GHRSST-PP data processing centre codes applicable to the GDS

The codes defined in Table A2.1 provide a unique identifier for each GHRSST-PP RDAC and GDAC data processing centre.

Table A2.1 GDS data centre prefix codes.

Prefix Code	Data centre name
MGDAC	US-GODAE Monterey GDAC
JGDAC	Jet Propulsion Laboratory GDAC
EUR	European RDAC
USA	United States RDAC
JAP	Japanese RDAC
SEASNET	SEASnet Tropical coverage RDAC
REMSS	Remote Sensing Systems, CA, USA
NASDA	National space Development Agency (of Japan)
ESA	European Space Agency
RSMAS	University of Miami, RSMAS.
TOHOKU	University of Tohoku, Japan
SOC	Southampton Oceanography Centre, UK
JPL	Jet Propulsion Laboratory
OSISAF	EUMETSAT Ocean and Sea Ice Satellite Applications Facility
AUST	Australian RDAC
USGODAE	US-GODAE
CORIOLIS	CORILOIS data centre, IFREMER, France.
MEDS	MEDS data centre, Ontario, Canada.

A2.2 GDS data set name codes

The codes defined in Table A2.2 provide a unique identifier for each data set referenced by the GDS.

Table A2.2 GDS Data set name codes (see Appendix A3 for a full description)

Input data stream	Description
ATS_NR__2P	ENVISAT AATSR near real time 1km SSTskin data
ATS_MET_2P	ENVISAT AATSR real time meteorological data product
AVHRR16_G	AVHRR NOAA-16 GAC derived SST data
AVHRR16_L	AVHRR NOAA-16 LAC derived SST data
AVHRR17_G	AVHRR NOAA-17 GAC derived SST data
AVHRR17_L	AVHRR NOAA-17 LAC derived SST data
SEVIRI	MSG-SEVIRI derived SST data
SEVIRI_SSI	MSG-SEVIRI derived SSI data
GOESE	GOES-E derived SST data
GOESE_SSI	GOES-E derived SSI data
GOESW	GOES-W derived SST data
GOESW_SSI	GOES-W derived SSI data
GMS	GMS derived SST data (GOES-9?)
GMS_SSI	GMS derived SSI data (GOES-9?)
AMSRE	AMSR-E derived SST data

AMSRE_WSP	AMSR-E derived wind speed data
AMSRE_WV	AMSR-E derived atmospheric water vapour data
TMI	TRMM TMI derived SST data
TMI_WSP	TRMM TMI derived wind speed data
TMI_WV	TRMM TMI derived atmospheric water vapour data
TMI_VIRS	TRMM TMI-VIIRS derived SST data
AIRS	EOS-AQUA AIRS derived SST
MODIS_A	EOS AQUA MODIS derived SST data
MODIS_T	EOS TERRA MODIS derived SST data
AMSR_ICE	AMSR derived ocean- ice data product
AMSRE_ICE	AMSR-E derived ocean ice data product
SSMIFxx_WSP	SSMI derived wind speed (xx refers to the DMSP satellite number)
SSTFND	SSTfnd gridded data file
SSTFNDUHR	Ultra-high resolution SSTfnd gridded data file

Appendix A3. GDS input data definitions

The following sub-sections collectively describe the data products that are used by the GDS.

A3.1 GDS Reference data sets

The GDS refers to the reference fields described in Table A3.1.1

Table A3.1.1 Reference data fields used by the GDS

ID	Name	Description	Data agreement	Reference
R1	10 day Minimum Climatological SST1m	Ten year (1985-1995) time series of daily 9km Pathfinder SST data. The reference SST is derived as consecutive 10 day minimum night time SST maps	None	Faugere et al. (2001)
R2	GDS n-day mean SSTfnd	The mean SSTfnd computed for a n- day period. This product is computed from IDSI-PMv1 SSTfnd data products in real time each day	N/A	GHRSST-PP GDS.
R3	DMSP SSM/I Daily and Monthly Polar Gridded Sea Ice Concentrations (NASA Team algorithm)	NSIDC SSM/I sea ice products in polar stereographic projection currently include DMSP-F8, F11 and F13 daily and monthly sea ice concentrations. Data, gridded at a resolution of 25 x 25 km, begin 25 June 1987. (NASA_TEAM algorithm) available in near real time.	None	Cavialeri et al (1992) ftp://sidas.colorado.edu/ftp/DATASETS/PASSIVE_MICROWAVE/POLAR_STEREO/DATA/SEAICE/SSMI/NASATEAM/
R4	MODIS/Terra Land Cover Type 96-Day L4 Global 1km ISIN Grid	Land Cover Classification product, MOD12Q1, identifies 17 classes of land cover in the International Geosphere-Biosphere Programme (IGBP) global vegetation classification scheme. The data set is based on 1 year of MODIS data 2000-2001 and includes a land water cover QC flag that provides a simple ocean water classification.	None	http://edcdaac.usgs.gov/modis/mod12q1.asp
R5	NOAA/NESDIS Aerosol Optical Thickness (AOT) 100km Product ID AERO100	Weekly 1 degree map of Aerosol Optical Thickness (AOT) based on a colmposite of AVHRR data	None	http://www.ssa.noaa.gov/cocoon/nsaa/products/search?datatype_family=AERO100&submit.x=22&submit.y=10
R6	OIv2: Reynolds Optimal Interpolated SST analysis v2	OI.v2 is a SST analysis produced by a blend of AVHRR and in situ data (ship and buoy). The analysis is produced operational weekly on a one-degree spatial grid.	None	Reynolds et al. (2002) http://www.emc.ncep.noaa.gov/research/cmb/sst_analysis/
R7	OI.v2 monthly climatology	monthly climatology produced from and other analyses with a base period of 1971-2000.	None	Xue et al (2003)) http://www.cpc.noaa.gov/products/predictions/30day/SSTs/s

				st_clim.html
R8	HadSST v1	Monthly 1 degree by 1 degree resolution SST and sea ice fields from HadISST1 from 1871 to date. Complete SST fields are reconstructed based on in situ measurements from ships and buoys through 1981 and from these blended with bias-corrected AVHRR SST from 1982 onwards		Rayner et al. (2003) http://www.metoffice.com/
R9	10 day Mean Climatological SST1m	Ten year (1985-1995) time series of daily 9km Pathfinder SST data. The reference SST is derived as consecutive 10 day mean night time SST maps	None	Faugere et al. (2001)
R10	Pathfider monthly SST climatology (Erosion filter version)	9.28km resolution monthly Pathfinder+Erosion Sea Surface Temperature climatology. 1985-1997	None	http://podaac.jpl.nasa.gov

A3.2 Satellite SST data streams used in the GDS

The following sections provide a reference to each of the satellite SST data streams that will be used in the GDS. Table A3.2.1 provides a summary of each data stream. Not all of these data will be used by all RDAC as some data are regional in coverage. Tables will be added as new data streams are brought on-line but tables will not be deleted if data streams are taken off-line as these may be used by the GHRSST-PP reanalysis project.

Table A3.2.1 Summary of satellite SST data streams used by the GDS

Code	GHRSST-PP Identification Code	Sensor name	Nadir FoV resolution	Coverage (space and time)	Variable	Data provider	Data Format	GHRSST-PP ingestion	Data Transport	Applicable Data Agreement
1	**ATS_NR__2P_SST**	AATSR	1km	Global with pseudo 3 day repeat	SSTskin	ESA	ATS_NR__2P (http://envisat.esa.int/dataproducts/aatsr/toc.htm)	EUR-RDAC GDAC	ftp pull from Southampton Oceanography Centre	ESA Cat-1 proposal ID1329 (Donlon et al)
2	**ATS_MET_2P_SST**	AATSR	10 arc min	Global, pseudo 3 day repeat	SSTskin	ESA	ATS_MET_2P (http://envisat.esa.int/dataproducts/aatsr/toc.htm)	EUR-RDAC GDAC	ftp from US-GODAE http://www.usgodae.org/cgi-bin/datalist.pl?generate=summary	ESA Cat-1 proposal ID1329 (Donlon et al)
3	**AVHRR16-G_SST**	AVHRR NOAA16	4 km (GAC)	Global, daily repeat	SST1m	NAVOCEANO	GHRSST-PP L2P	GDAC and all RDAC	ftp from US-GODAE http://www.usgodae.org/cgi-bin/datalist.pl?generate=summary	None
4	**AVHRR16-L_SST**	AVHRR NOAA16	1-2 km (LAC)	Regional, at least daily repeat, day and night	SST1m	NAVOCEANO	GHRSST-PP L2P	GDAC and all RDAC	ftp from US-GODAE http://www.usgodae.org/cgi-bin/datalist.pl?generate=summary	None
5	**AVHRR16-L_SST**	AVHRR NOAA16	2 km (LAC)	Regional, at least daily repeat, day and night	SSTsubskin	EUMESAT O&SI SAF	http://www.meteorologie.eu.org/safo	EUR RDAC	ftp from EUMESAT O&SI SAF ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI	None
6	**AVHRR17-G_SST**	AVHRR NOAA17	4 km (GAC)	Global, daily repeat	SST1m	NAVOCEANO	GHRSST-PP L2P	GDAC and all RDAC	ftp from US-GODAE http://www.usgodae.org/cgi-bin/datalist.pl?generate=summary	NAVOCEANO
7	**AVHRR17-L_SST**	AVHRR NOAA17	1-2 km (LAC)	Regional, at least daily repeat, day and night	SST1mn	NAVOCEANO/	GHRSST-PP L2P	GDAC and all RDAC	ftp from US-GODAE http://www.usgodae.org/cgi-bin/datalist.pl?generate=summary	NAVOCEANO
8	**AVHRR17-L_SST**	AVHRR NOAA17	1-2 km (LAC)	Regional, at least daily repeat, day and night	SSTsubskin	EUMESAT O&SI SAF	http://www.meteorologie.eu.org/safo	EUR RDAC	ftp from EUMESAT O&SI SAF ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI	None
9	**SEVIRI_SST**	MSG SEVIRI	10 km	Regional, 3 hour repeat	SSTsubskin	EUMETSAT O&SI SAF	http://www.meteorologie.eu.org/safo/gb_html_sst/regional/regsst_manual.pdf	EUR RDAC	ftp from EUMESAT O&SI SAF ftp://ftp.ifremer.fr/pub/ifremer	None

							http://www.meteorologie.eu.org/safo/gb_html_sst/atlantic/atlsst_manual.pdf		er/cersat/SAFOSI	
10	**GOES-12SAF_SST**	GOES-12	4 km	Regional, 3 hour repeat	SST1m	EUMETSAT O&SI SAF	http://www.meteorologie.eu.org/safo/gb_html_sst/regional/regsst_manual.pdf http://www.meteorologie.eu.org/safo/gb_html_sst/atlantic/atlsst_manual.pdf	EUR RDAC	ftp from EUMESAT O&SI SAF ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI	None
11	**GOES-10_SST**	GOES-10	6 km	Regional, ½ hour repeat	SST1m	NAVOCEANO & PO.DAAC	http://podaac.jpl.nasa.gov/noaa_goes	JPL GDAC	ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/goes/NOAA	None
12	**GOES-12_SST**	GOES-12	6 km	Regional, 1/2 hour repeat	SST1m	NAVOCEANO & PO.DAAC	http://podaac.jpl.nasa.gov/noaa_goes	JPL GDAC	ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/goes/NOAA	None
13	**GOES-10NAV_SST**	GOES-10	4 km	Regional, ½ hour repeat	SST1m	NAVOCEANO	http://www.us-godae.org	GDAC	ftp://www.us-godae.org	None
14	**AMSRE_SST**	AMSR-E	25 km grid	Global, daily repeat	SSTsubskin	REMSS and L2P	http://www.remss.com/amsr/amsr_data_description.html and GHRSST-PP L2P	JPL GDAC and all RDAC	ftp from REMSS ftp.ssmi.com	Registration required
15	**TMI_SST**	TRMM-TMI	25 km grid	Global, daily repeat	SSTsubskin	REMSS and L2P	http://www.remss.com/tmi/tmi_browse.html and GHRSST-PP L2P	JPL GDAC and all RDAC	ftp from REMSS ftp.ssmi.com	None
16	**TMI-VIRS_SST**	TRMM-VIRS	2km	Global, daily repeat	SSTsubskin	JAXA	TBD	GDAC and all RDAC	TBD	TBD
17	**MODISA_SST**	MODIS AQUA	1 km	Global	SSTskin	JPL PO.DAAC	http://www.ghrsst.pp.org/documents	JPL GDAC and all RDAC	ftp from ftp.podaac.jpl.nasa.gov expected summer 2004	none
~~**18**~~	~~**MODIS-T**~~	~~MODIS TERRA~~	~~1 km or 4km TBC~~	~~TBD but potentially global~~	~~SSTdepth~~	~~In Progress~~	~~TBD~~	~~TBD~~	~~Real time service not currently available for the GDS but expected in the GDSv2~~	~~TBD~~
19	**WindSAT_SST**	WindSAT	25 km grid	Global, daily, repeat	SSTsubskin	REMSS	GHRSST-PP L2P – not available yet	GDAC and all RDAC	ftp from REMSS ftp.ssmi.com	Unknown
20	**GOES-9_SST**	VISSR	4 km	Regional	SSTskin/SSTsubskin	TBD	TBD	Japan RDAC	TBD	TBD
21	**MTSAT-1R_SST**	S-VISSR	5 km	Regional	SSTskin/SSTsubskin	JMA	Not available until 2004	JAPAN RDAC	TBD	TBD
22	**AIRS_SST**	AIRS	45 km	Global	SSTskin	NASA	TBD	JPLGDAC	TBA	

A3.2.1 Advanced Along Track Scanning Radiometer (AATSR) GHRSST-PP data streams.

GHRSST-PP ID	**ATS_NR__2P_SST:** high resolution(1.1km) SSTskin data product **ATS_MET_2P_SST:** low resolution (10 arc minute grid) averaged SSTskin and BT data product
Point of contact	Craig Donlon (craig.donlon@metoffice.com +44 01392 886622)
Data provider	European Space Agency (ESA)
Data transport	ftp pull service from GHRSST/ESA server
Data agreement	AATSR data are provided to the GHRSST-PP Science Team and other researchers free of charge within the framework of the GHRSST-PP under an ESA Category-1 data agreement ID is 1329 (see http://projects.esa-ao.org => ESA Data Policy=> Terms and conditions). Under this agreement, ESA make available in real time, to PIs named of Cat-1 1329. Additional researches can be added to this agreement by sending a signed copy of the ESA Category-1 data agreement and a short summary of the work that will be carried out using the AATSR data within the GHRSST-PP project.
Data products provided to GHRSST-PP	**ATS_NR__2P_SST:** Gridded Surface Temperature (GST) Product at full resolution (1.1km). **ATS_MET_2P_SST:** The Meteo product is a fast delivery product designed for use by meteorological offices, and contains averaged Brightness Temperatures and SST at 10 arc minute resolution. This is available in either ENVISAT format or BUFR format.
Data format document	A full specification of ENVISAT AATSR ATS_NR__2P and ATS_MET_2P data products can be found at http://envisat.esa.org/aatsr/. BUFR format http://envisat.esa.int/dataproducts/aatsr/CNTR6-1-2.htm#eph.aatsr.aatsrdf.2p.ATS_MET_2P
Primary GHRSST-PP entry point	EUR-RDAC (Medspiration Project), ESA ftp server for GHRSST-PP (METEO products only)
Secondary GHRSST-PP entry point	GDAC via ftp pull
Notes	Initially, only METEO data products will be available by ftp service.

A3.2.2 Advanced Very High Resolution Radiometer (AVHRR) GHRSST-PP data streams.

GHRSST-PP ID	**AVHRR16-G_SST**: GAC format NOAA 16 AVHRR data **AVHRR16-L_SST**: LAC format NOAA 16 AVHRR data **AVHRR17-G_SST**: GAC format NOAA 17 AVHRR data **AVHRR17-L_SST**: LAC format NOAA 17 AVHRR data
Points of contact	1. Jorge Vasquez (jv@pacific.jpl.nasa.gov) Doug May (mayd@navocean.navy.mil) 2. Jorge Vasquez (jv@pacific.jpl.nasa.gov) Doug May (mayd@navocean.navy.mil) 3. Pierre LeBorgne (leborgne@meteo.fr) 4. Michel Petit (Michel.Petit@ird.fr)
Data provider	1. PO.DAAC and US-Navy 2. PO.DAAC and US-Navy 3. EUMETSAT O&SI SAF 4. Survey of Environment Assisted by Satellites (SEASnet) project
Data transport	1. L2-LAC 2.2 km MCSST (Regional coverage) ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/avhrr/navoceano_hrpt_lac/data/L2 2. Orbital ungridded 9km MCSST

	ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/avhrr/navoceano_mcsst/ 3. ftp pull from EUMETSAT O&SI SAF server ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI 4. ftp
Data agreement	1. None 2. None 3. None 4. TBD
Data products provided to GHRSST-PP	1. 2.2 km MCSST product using both Local Area Coverage (LAC) and High Resolution Picture Transmission (HRPT) SST calculations. Coverage includes the East and West Coasts of the United States, the Mediterranean, and parts of the Indian Ocean and the Australian Coasts. 2. Single orbit ungridded 9km Multi-channel sea-surface temperatures (MCSST). Improved cloud retrieval algorithms applied by NAVOCEANO have significantly increased the number of SST pixels flagged as cloud free, as compared to previous MCSST data sets. 3. 3 hourly SSTsubskin (bias corrected using night time buoy data) stereo polar grid 4km resolution over the European Seas 4. SEASnet intend to produce GHRSST-PP L2P and L2C data products using LAC AVHRR SST in the tropical areas (Receiving stations: Reunion Island, French Guyana, Canary Islands, New Caledonia). Gridded data on a 0.01° angle. Coverage over LAC receiving stations (see ?? for image).
Data format document	1. http://podaac.jpl.nasa.gov/pub/sea_surface_temperature/avhrr/navoceano_mcsst/doc/avhrr_navoceano.html 2. http://podaac.jpl.nasa.gov/order/order_sstemp.html#Product143 3. Product manual at http://www.meteorlogie.eu.org/safo 4. netCDF (GHRSST-PP data format (see section 2.7)
Primary GHRSST-PP entry point	1. JPLGDAC 2. EUR RDAC 3. SEASnet RDAC 4. JAP RDAC 5. AUST RDAC
Secondary GHRSST-PP entry point	
Notes	1. The AVHRR sensor is used at all RDAC centres and contacts are given for each centre. 2. Additional LAC coverage data available at RDAC via local receiving stations

A3.2.3 METEOSAT Second Generation Spinning Enhanced Visible and Infrared Imager (SEVIRI) GHRSST-PP data streams.

GHRSST-PP ID	**SEVIRI_SST**
Point of contact	Pierre LeBorgne (leborgne@meteo.fr)
Data provider	EUMETSAT O&SI SAF
Data transport	ftp pull from EUMETSAT O&SI SAF server ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI
Data agreement	None for EUMETSAT,
Data products provided to GHRSST-PP	Gridded SSTsubskin over the Atlantic Ocean from 100°W – 45°E and 60°N – 90°N at a grid resolution of 0.1°

Data format document	http://www.meteorologie.eu.org/safo/gb_html_sst/regional/regsst_manual.pdf http://www.meteorologie.eu.org/safo/gb_html_sst/atlantic/atlsst_manual.pdf
Primary GHRSST-PP entry point	EUR RDAC
Secondary GHRSST-PP entry point	
Notes	Should be available by spring 2004

A3.2.4 Geostationary Operational Environmental Satellite (GOES) 10 GHRSST-PP data streams.

GHRSST-PP ID	**GOES10_SST**
Point of contact	Jorge Vasquez (Jorge.vasquez@podaac.jpl.nasa.gov)
Data provider	NOAA/NESDIS
Data transport	ftp from ftp://podaac.jpl.nasa.gov/pub/sea_surface_temperature/goes/NOAA/
Data agreement	None
Data products provided to GHRSST-PP	GOES L4 6km Near-Real-Time SST (NOAA/NESDIS) Sea Surface Temperatures (SSTs) are derived from the series of Geostationary Environmental Satellites (GOES). SSTs are calculated using a regression onto the channel brightness temperatures on board the GOES instrument.
Data format document	http://podaac.jpl.nasa.gov/cgi-bin/dcatalog/fam_summary.pl?sst+goes
Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	
Notes	

A3.2.5 Geostationary Operational Environmental Satellite (GOES) 12 GHRSST-PP data streams.

GHRSST-PP ID	**GOES12_SST**
Point of contact	1. Jorge Vasquez (Jorge.vasquez@podaac.jpl.nasa.gov) 2. Pierre LeBorgne (leborgne@meteo.fr)
Data provider	1. PO.DAAC at JPL 2. EUMETSAT O&SI SAF
Data transport	1. ftp pull from ftp.po.daac.jpl.nasa.gov 2. ftp pull from EUMETSAT O&SI SAF server ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI
Data agreement	1. None 2. None
Data products provided to GHRSST-PP	1. GOES L4 6km Near-Real-Time SST (NOAA/NESDIS) 2. Gridded Surface Solar Irradiance (SSI) over the Atlantic Ocean from 100°W – 45°E and 60°N – 90°N at a grid resolution of 0.1° Data are available in real time at 3 hourly intervals.
Data format document	1. http://podaac.jpl.nasa.gov 2. http://www.meteorologie.eu.org/safo/gb_html_ssi/atlantic/ssi_manual.pdf
Primary GHRSST-PP entry point	1. JPL GDAC 2. EUR RDAC
Secondary GHRSST-PP entry	

point	
Notes	

A3.2.6 Advanced Microwave Scanning Radiometer-E (AMSR-E) GHRSST-PP data streams.

GHRSST-PP ID	**AMSRE_SST**
Point of contact	Chelle Gentemann (gentemann@remss.com)
Data provider	Remote Sensing systems (REMSS)
Data transport	ftp pull from ftp.ssmi.com/amsr
Data agreement	Agreement with GHRSST-PP Science Team for use of data products
Data products provided to GHRSST-PP	Each binary data file consists of twelve daily 0.25 x 0.25 degree grid (1440,720) byte maps. For daily files, six daytime maps in the following order, Time (UTC), Sea Surface Temperature (SST), 10 meter Surface Wind Speed (WSPD), Atmospheric Water Vapour (VAPOR), Cloud Liquid Water (CLOUD), and Precipitation Rate (RAIN), are followed by six night time maps in the same order.
Data format document	http://www.remss.com/amsr/amsr_data_description.html
Primary GHRSST-PP entry point	GDAC
Secondary GHRSST-PP entry point	RDAC
Notes	Data files are updated in real time as orbital data arrives at RSS and is processed. RDAC are requested to note the filename creation date and time as an indicator of new data.

A3.2.7 TRMM Microwave Imager (TMI) GHRSST-PP data streams.

GHRSST-PP ID	**TMI_SST**
Point of contact	Chelle Gentemann (gentemann@remss.com)
Data provider	Remote Sensing Systems (REMSS)
Data transport	ftp pull from ftp.ssmi.com/tmi
Data agreement	None
Data products provided to GHRSST-PP	Daily binary data files consists of fourteen 0.25 x 0.25 degree grid (1440,320) byte maps. Seven ascending maps in the following order: Time (T), Sea Surface Temperature (SST), 10-meter Surface Wind Speed using 11 GHz (Z), 10-meter Surface Wind Speed using 37 GHz (W), Atmospheric Water Vapour (V), Cloud Liquid Water (L), and Precipitation Rate (R), are followed by seven descending maps in the same order.
Data format document	http://www.remss.com/tmi/tmi_description.html
Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	EU RDAC, JAP RDAC
Notes	TMI data are also available from NSADA Data files are updated in real time as orbital data arrives at RSS and is processed. The filename containing incomplete coverage is denoted with the extension .tm. The .tm extension is removed when all data for the given data have been included in the data file. RDAC are requested to note the filename creation date and time as an indicator of new data.

3.2.8 TRMM Visible and Infrared Scanner (VIRS) GHRSST-PP data streams.

GHRSST-PP ID	**VIRS_SST**
Point of contact	Hiroshi Kawamura (kamu@ ocean.caos.tohoku.ac.jp)
Data provider	NASDA currently in discussion
Data transport	[TBD]
Data agreement	[TBD]
Data products provided to GHRSST-PP	VIRS SST spanning ~40ºN – 40ºS spatial resolution of ~2 km at nadir. Contemporaneous with TMI data and extremely useful for understanding differences between IR and MW SST measurements
Data format document	[TBD]
Primary GHRSST-PP entry point	Japanese RDAC
Secondary GHRSST-PP entry point	Other RDAC/GDAC
Notes	TRMM VIRS data are currently in discussion

A3.2.9 MODIS GHRSST-PP data streams.

GHRSST-PP ID	**MODISA_SST**
Point of contact	Ed Armstrong (ed@seastar.jpl.nasa.gov)
Data provider	PO.DAAC, JPL
Data transport	ftp from ftp://podaac.jpl.nasa.gov
Data agreement	None
Data products provided to GHRSST-PP	GHRSST-PP L2P data format MODIS AQUA SST.
Data format document	GHRSST_PP Data processing specification (GDS)
Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	Other RDAC
Notes	L2P data products expected summer 2004

A3.2.10 WindSAT GHRSST-PP data streams.

GHRSST-PP ID	**Windsat_SST**
Point of contact	Chelle Gentemann (gentemann@remss.com)
Data provider	Remote Sensing Systems (REMSS)
Data transport	TBD
Data agreement	Unknown (To be determined)
Data products provided to GHRSST-PP	Expected: Daily binary data files consists of fourteen 0.25 x 0.25 degree grid (1440,320) byte maps. Seven ascending maps in the following order: Time (T), 10-meter Surface Wind Speed using 11 GHz (Z), 10-meter Surface Wind Speed using 37 GHz (W), Atmospheric Water Vapour (V), Cloud Liquid Water (L), and Precipitation Rate (R), are followed by seven descending maps in the same order.
Data format document	Not available at this time

Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	EU RDAC, JAP RDAC

A3.2.11 GOES-9 (Japanese area) GHRSST-PP data streams.

GHRSST-PP ID	**GOES9_SST**
Point of contact	TBD
Data provider	Japan Meteorological Agency
Data transport	ftp pull from TBD
Data agreement	Unknown (To be determined)
Data products provided to GHRSST-PP	SST @ 5km spatial resolution (nadir) and 1 hour repeat located over Japan derived from the Stretched Visible Infrared Spin Scan Radiometer (S-VISSR). See http://www.jma.go.jp for more information.
Data format document	Not available at this time
Primary GHRSST-PP entry point	Japan RDAC
Secondary GHRSST-PP entry point	None

A3.2.12 MTSAT GHRSST-PP data streams.

GHRSST-PP ID	**MTSAT-SST**
Point of contact	TBD
Data provider	Japan Meteorological Agency
Data transport	ftp pull from TBD
Data agreement	Unknown (To be determined)
Data products provided to GHRSST-PP	TBD
Data format document	Not available at this time (January 2004)
Primary GHRSST-PP entry point	Japan RDAC
Secondary GHRSST-PP entry point	None

A3.2.13 AIRS GHRSST-PP data streams.

GHRSST-PP ID	**AIRS_SST**
Point of contact	D. Hagan d.hagan@jpl.nasa.gov
Data provider	NASA JPL
Data transport	ftp pull from JPL [TBD]
Data agreement	NASA-GHRSST-PP
Data products provided to GHRSST-PP	AIRS GHRSST-PP format L2 data file stripped from the AIRS L2 standard data products containing: • SST • Time of SST • Geolocation (latitude, longitude) • Ret QC Flag (AIRS retrieval confidence value) • Low/high cloud estimate (Ci identification, Low cloud (St))

	• Totals water vapour • Land flag • Rain flag 45 km resolution, Daily global coverage
Data format document	http://daac.gsfc.nasa.gov/atmodyn/airs
Primary GHRSST-PP entry point	GDAC (JPL)
Secondary GHRSST-PP entry point	None
Notes	L2/L2P AIRS-GHRSST-PP format data should be available in late 2004.

A3.3 GDS auxiliary satellite data streams

Table 3.3.1 provides a summary of the auxiliary satellite data streams that will be used in the GDS.

Table 3.3.1 Summary of input satellite auxiliary data streams considered by the GDS.

Code	GHRSST-PP Identification Code	Sensor name	Nadir FoV resolution	Coverage (space and time)	Variable	Data provider	Data Format	GHRSST-PP ingestion	Data Transport	Applicable Data Agreement
101	**NISE**	SSM/I	25km	Global, daily	NSIDC SSM/ Sea Ice data Sea Ice and extent	NSIDC	http://nsidc.org/data/docs/daac/nise1_nise.gd.html	JPL GDAC	ftp://sidads.colorado.edu/DATASETS/PASSIVE_MICROWAVE/POLAR_STEREO/DATA/SEA_ICE/SSMI/NASATEAM/F13/NEAR_REAL_TIME	Registration required at http://www.nsidc.org
102	**AMSRE-ICE**	AMSR-E	25 km (6km also available)	Daily	Brightness temperature, Sea Ice concentration & snow depth on polar grids	NSIDC	http://www.nsidc.org/data/ae_si12.html	RDAC & GDAC	ftp pull from ftp://sidads.colorado.edu	regisdtration and password are required at http://www.nsidc.org
103	**SEVIRI-SSI**	MSG SEVIRI	0.1 x 0.1 degree	3 hourly	SSI	EUMETSAT O&SI SAF	http://www.meteorologie.eu.org/safo/gb_html_ssi/atlantic/ssi_manual.pdf	EUR RDAC	ftp pull from EUMETSAT O&SI SAF server ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI	None
104	**GOES10-SSI**						TBD			
105	**GOES12-SSI**						TBD			
107	**AMSRE-WSP**	AMSR-E	0.25° grid	Global	Gridded surface (10m) wind speed	REMSS	http://www.remss.com/amsr/amsr_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com	Registration required (brewer@remss.com)
108	**AMSRE-WV**	AMSR-E	0.25° grid	Global	Gridded atmospheric water vapour	REMSS	http://www.remss.com/amsr/amsr_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com	Registration required (brewer@remss.com)
109	**AMSRE-CLD**	AMSR-E	0.25° grid	Global	Gridded cloud liquid water vapour	REMSS	http://www.remss.com/amsr/amsr_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com	Registration required (brewer@remss.com)
113	**TMI-WSP**	TRMM TMI	0.25° grid	40°N – 40°S, 180E -180W	Gridded surface (10m) wind speed	REMSS	http://www.remss.com/tmi/tmi_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com/tmi	none

114	**TMI-WV**	TRMM TMI	0.25° grid	40°N – 40°S, 180E -180W	Gridded atmospheric water vapour	REMSS	http://www.remss.com/tmi/tmi_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com/tmi	none
115	**TMI-CLD**	TRMM TMI	0.25° grid	40°N – 40°S, 180E -180W	Cloud liquid water vapour	REMSS	http://www.remss.com/tmi/tmi_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com/tmi	None
116	**SSMI-WSP**	DMSP SSM/I	0.25° grid	Global	Gridded surface (10m) wind speed	REMMS	http://www.remss.com/ssmi/ssm_data_description.html	JPL GDAC	ftp pull ftp.ssmi.com/ssmi	None
117	**Windsat_WSP**	Windsat	0.25° grid	Global	Gridded surface (10m) wind speed	TBC	TBC	TBC	TBC	TBC
118	**AVHRR-AOD**	AVHRR	100km grid	Global	Weekly Gridded atmospheric optical depth (AOD)	NOAA	http://www.osdpd.noaa.gov/PSB/EPS/Aerosol/Aerosol.html	RDAC and GDAC	http://www.saa.noaa.gov/cocoon/nsaa/products (LAS type server)	None

A3.3.1 National Snow and Ice Data Centre SSMI/I sea Ice products GHRSST-PP data streams.

GHRSST-PP ID	**NISE**
Point of contact	Donald Cavalieri (Donald.J.Cavalieri@nasa.gov)
Data provider	National Snow and ice data Center (NSIDC, http://nsidc.org)
Data transport	ftp pull from ftp://sidads.colorado.edu/DATASETS/PASSIVE_MICROWAVE/POLAR_STEREO/DATA/SEA_ICE/SSMI/NASATEAM/F13/NEAR_REAL_TIME
Data agreement	Registration required at NSIDC
Data products provided to GHRSST-PP	Near Real-Time SSM/I EASE-Grid Daily Global Ice Concentration and Snow Extent product (Near real-time Ice and Snow Extent, NISE) provides daily, global near real-time maps of sea ice concentrations and snow extent. The National Snow and Ice Data Center (NSIDC) created the NISE product using passive microwave data from the Defence Meteorological Satellite Program (DMSP) F13 Special Sensor Microwave/Imager (SSM/I). Sea ice concentration and snow extent maps are provided in two 25 km azimuthal, equal-area projections: the Southern Hemisphere 25 km low resolution (SI) and Northern Hemisphere 25 km low resolution (NI) Equal-Area Scalable Earth-Grids (EASE-Grids). Data in Hierarchical Data Format - Earth Observing System (HDF-EOS) format, and browse files in GIF and HDF formats, are updated daily and are available via ftp for two weeks after initial posting.
Data format document	http://nsidc.org/data/docs/daac/nise1_nise.gd.html
Primary GHRSST-PP entry point	JPL GDAC/RDAC
Secondary GHRSST-PP entry point	none
Notes	This is a secondary data stream as the AMSR-E 12.5 km sea ice data products are preferred. However these are not finalised or available as of Feb 2003

A3.3.2 AMSR-E Sea Ice GHRSST-PP data streams.

GHRSST-PP ID	**AMSRE-ICE**
Point of contact	D. Cavalieri (don@cavliere.gsfc.nasa.gov)
Data provider	RemNAtional Snow and Ice Data Centre (NSIDC)
Data transport	ftp pull from ftp://sidads.colorado.edu
Data agreement	Registration required at NSIDC (no cost)
Data products provided to GHRSST-PP	The Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) instrument on the NASA EOS Aqua satellite provides global passive microwave measurements of land, ocean, and atmospheric variables for the investigation of water and energy cycles. This Level-3 gridded product (AE_SI12) includes brightness temperatures (TBs) at 18.7 through 89.0 GHz, sea ice concentration, and snow depth over sea ice. Data are mapped to a polar stereographic grid at 12.5 km spatial resolution. Sea ice concentrations and TBs include daily ascending averages, daily descending averages, and daily averages; snow depth over sea ice data is a five-day running average. Data are stored in HDF-EOS format and are available via FTP, CD-ROM, 8-mm tape, DLT, or DVD-ROM.

	Data Citation The following example shows how to cite the use of this data set in a publication. List the principal investigators, year of data set release (2004), data set title and version, date of the version you used, publishers (NSIDC), and digital media. **Example citation:** Cavalieri, D., and J. Comiso. 2004. AMSR-E/Aqua Daily L3 12.5 km Tb, Sea Ice Conc., & Snow Depth Polar Grids V001, March 2004. Boulder, CO, USA: National Snow and Ice Data Center. Digital media.
Data format document	http://www.nsidc.org/dcata/ae_si12.html
Primary GHRSST-PP entry point	JPL GDAC/RDAC
Secondary GHRSST-PP entry point	none
Notes	Data files are available on a dialy basis. A 6km spatial resolution data product is also available.

A3.3.3 METEOSAT Second Generation Spinning Enhanced Visible and Infrared Imager (SEVIRI) GHRSST-PP data streams.

GHRSST-PP ID	**SEVIRI-SSI**
Point of contact	Pierre LeBorgne (leborgne@meteo.fr)
Data provider	EUMETSAT O&SI SAF
Data transport	ftp pull from EUMETSAT O&SI SAF server ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI
Data agreement	None for EUMETSAT,
Data products provided to GHRSST-PP	Gridded Surface Solar Irradiance (SSI) over the Atlantic Ocean from 100°W – 45°E and 60°N – 90°N at a grid resolution of 0.1°
Data format document	Product manual at http://www.meteorlogie.eu.org/safo
Primary GHRSST-PP entry point	EUR RDAC
Secondary GHRSST-PP entry point	US-GODAE via IFREMER ftp push
Notes	

A3.3.4 Geostationary Operational Environmental Satellite (GOES) 10 GHRSST-PP data streams.

GHRSST-PP ID	**GOES10-SSI**
Point of contact	Eileen Maturi (Eileen.mauturi@noaa.gov)
Data provider	NOAA/NESDIS
Data transport	ftp from JPL – currently in preparation
Data agreement	None
Data products provided to GHRSST-PP	[TBD]
Data format document	[TBD]

Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	
Notes	

A3.3.5 Geostationary Operational Environmental Satellite (GOES) 12 GHRSST-PP data streams.

GHRSST-PP ID	**GOES12_SSI**
Point of contact	1. Eileen Maturi (Eileen.mauturi@noaa.gov) 2. Pierre LeBorgne (leborgne@meteo.fr)
Data provider	1. NOAA/NESDIS 2. EUMETSAT O&SI SAF
Data transport	1. ftp pull from JPL GDAC 2. ftp pull from EUMETSAT O&SI SAF server ftp://ftp.ifremer.fr/pub/ifremer/cersat/SAFOSI
Data agreement	1. [TBD] 2. None
Data products provided to GHRSST-PP	1. [TBD] 2. Gridded Solar Irradiance (SSI) over the Atlantic Ocean from 100°W – 45°E and 60°N – 90°N at a grid resolution of 0.1°
Data format document	1. [TBD] 2. http://www.meteorologie.eu.org/safo/gb_html_ssi/atlantic/ssi_manual.pdf
Primary GHRSST-PP entry point	1. JPL GDAC 2. EUR RDAC
Secondary GHRSST-PP entry point	
Notes	

A3.3.6 Advanced Microwave Scanning Radiometer-E (AMSR-E) GHRSST-PP data streams.

GHRSST-PP ID	**AMSRE-WSP**
Point of contact	Chelle Gentemann (gentemann@remss.com)
Data provider	Remote Sensing systems (REMSS)
Data transport	ftp pull from ftp.ssmi.com/amsr
Data agreement	Agreement with GHRSST-PP Science Team for use of data products
Data products provided to GHRSST-PP	Each binary data file consists of twelve daily 0.25 x 0.25 degree grid (1440,720) byte maps. For daily files, six daytime maps in the following order, Time (UTC), 10 meter Surface Wind Speed (WSPD), Atmospheric Water Vapour (VAPOR), Cloud Liquid Water (CLOUD), and Precipitation Rate (RAIN), are followed by six night time maps in the same order.
Data format document	http://www.remss.com/amsr/amsr_data_description.html
Primary GHRSST-PP entry point	GDAC
Secondary GHRSST-PP entry point	RDAC
Notes	

A3.3.7 TRMM Microwave Imager (TMI) GHRSST-PP data streams.

GHRSST-PP ID	**TMI-WSP**
Point of contact	Chelle Gentemann (gentemann@remss.com)
Data provider	Remote Sensing Systems (REMSS)
Data transport	ftp pull from ftp.ssmi.com/tmi
Data agreement	None
Data products provided to GHRSST-PP	Binary data files consists of fourteen 0.25 x 0.25 degree grid (1440,320) byte maps. Seven ascending maps in the following order: Time (T), 10-meter Surface Wind Speed using 11 GHz (Z), 10-meter Surface Wind Speed using 37 GHz (W), Atmospheric Water Vapour (V), Cloud Liquid Water (L), and Precipitation Rate (R), are followed by seven descending maps in the same order. Data files are updated in real time as orbital data arrives at RSS and is processed. The filename containing incomplete coverage is denoted with the extension .tm. The .tm extension is removed when all data for the given data have been included in the data file. RDAC are requested to note the filename creation date and time as an indicator of new data.
Data format document	http://www.remss.com/tmi/tmi_description.html
Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	EU RDAC, JAP RDAC
Notes	1. TMI data only cover 40°N – 40°S in low earth orbit configuration 2. TMI data are also available from NSADA

A3.3.8 SSM/I surface (10m) wind speed GHRSST-PP data streams.

GHRSST-PP ID	**SSMI-WSP**
Point of contact	D. Smith (smith@remss.com)
Data provider	Remote Sensing Systems (REMSS)
Data transport	ftp pull from ftp.ssmi.com/tmi
Data agreement	None
Data products provided to GHRSST-PP	Daily binary data files consists of fourteen 0.25 x 0.25 degree grid 0.25 x 0.25 degree grid (1440,720) byte maps. For daily files, five morning maps in the following order, Time (T), 10 meter Surface Wind Speed (W), Atmospheric Water Vapour (V), Cloud Liquid Water (L), and Precipitation Rate (R)
Data format document	http://www.remss.com/ssmi/ssmi_description.html
Primary GHRSST-PP entry point	JPL GDAC
Secondary GHRSST-PP entry point	EU RDAC, JAP RDAC
Notes	

A3.3.9 Windsat GHRSST-PP data streams.

GHRSST-PP ID	**WINDSAT-WSP**
Point of contact	Chelle Gentemann (gentemann@remss.com)
Data provider	Remote Sensing systems (REMSS)

Data transport	TBD
Data agreement	Agreement with GHRSST-PP Science Team for use of data products
Data products provided to GHRSST-PP	TBD
Data format document	TBD
Primary GHRSST-PP entry point	GDAC and RDAC
Secondary GHRSST-PP entry point	RDAC
Notes	Data available in 2004

A3.3.10 Aerosol optical thickness/depth (AOD)

GHRSST-PP ID	**AVHRR-AOD**
Point of contact	TBC
Data provider	NOAA
Data transport	LAS server at http://www.saa.noaa.gov/cocoon/nsaa/products/search?datatype_family=AERO100&submit.x=21&submit.y=15
Data agreement	None
Data products provided to GHRSST-PP	NOAA product Current retrievals of aerosol optical depths (AOD) are derived from AVHRR/3 reflectance channels 1 (0.63 um), 2 (0.83 um), and possibly 3A (1.61 um), from the two NOAA platforms: NOAA-16 (local equatorial crossing time, EXT~2 PM) and NOAA-17 (EXT~10 AM). Data are available as weekly files of 100km gridded fields. GIF images are available on a daily basis
Data format document	http://www.osdpd.noaa.gov/PSB/EPS/Aerosol/Aerosol.html
Primary GHRSST-PP entry point	GDAC
Secondary GHRSST-PP entry point	RDAC
Notes	Daily analysis products are experimental and only image files are currently available (Jan 2004)

A3.4 GDS in situ data streams

Table 3.4.1 provides a summary of the in situ data that are available to the GHRSST-PP in near real time (2-3 day delay).

Table 3.4.1 Summary of input in situ data streams considered by the GDS.

GHRSST-PP Identification Code	Platform	Variables	Data provider	Data Format	GHRSST-PP ingestion	Data Transport	Applicable Data Agreement
TAO	Tropical Pacific Moored buoy array (TAO)	SST1m, SSTdepth, + others$	PMEL	http://www.pmel.noaa.gov/tao/index.shtml	GAC and RDAC	GTS	
PIRATA	Tropical Atlantic Moored buoy	SST1m, SSTdepth, + others$	PMEL/AOML/IFREMER	http://www.pmel.noaa.gov/pirata/	GAC and RDAC	GTS	
ARGO	Profiling float	SSTdepth	CORIOLIS data centre (IFREMER)	http://www.ifremer.fr/coriolis/cdc/default.htm		GDS and ftp	None
RADIOM	Radiometer	SSTskin	None specific	Various	RDAC	ftp	Permission
CORIOLIS-SHIP CORILIS-PROFILE	Ships, profiles, buoys	Various	CORIOLIS data centre (IFREMER)	http://www.ifremer.fr/coriolis/cdc/default.htm	EUR RDAC	GTS and ftp	None
MEDS-DRIFT MEDS-MOOR MEDS-PROFILE	Drifting buoy	SST1m	Fisheries and Oceans service, Canada	http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Home_e.htm	GDAC/RDAC	GTS and ftp	None
JMA	Ships, profiles, buoys	SST1m	Japanese Meteorological Agency	TBC	RDAC	TBC	TBC
MAWS	Met Office, UK buoys	SST1m	Met Office, UK	TBC	RDAC	TBC	TBC
FNMOC	Ships, profiles, buoys	SST1m	Fleet Numerical Meteorological and Oceanography Centre FNMOC) via US-GODAE	http://www.usgodae.fnmoc.navy.mil	RDAC	ftp pull	None

$ wind speed, air temperature, atmospheric surface pressure, short-wave radiation

A3.4.1 Tropical Atmosphere Ocean (TAO) buoys GHRSST-PP data streams.

GHRSST-PP ID	**TAO**
Point of contact	atlasrt@noaa.gov
Data provider	Pacific Marine Environmental Research Laboratory (PMEL)
Data transport	GTS see http://www.wmo.ch/web/www/TEM/gts.html or by ftp from PMEL see http://www.pmel.noaa.gov/tao/proj_over/availability.html for availability
Data agreement	None
Data products provided to GHRSST-PP	Wind speed, wind direction, Air temperature, relative humidity, rainfall, Short wave radiation, long wave radiation, air pressure, SST at various depths, salinity see http://www.pmel.noaa.gov/tao/proj_over/sensors.shtml for full details of sensors and available data.
Data format document	ASCII or netCDF data files available on demand
Primary GHRSST-PP entry point	TBD
Secondary GHRSST-PP entry point	
Notes	

A3.4.2 Pilot Research Moored Array in the Tropical Atlantic (PIRATA) buoys GHRSST-PP data streams.

GHRSST-PP ID	**PIRATA**
Point of contact	Jaques Servain (jaques.servain@ird.fr) http://www.brest.ird.fr/pirata/piratafr.html
Data provider	Pacific Marine Environmental Research Laboratory (PMEL)
Data transport	GTS see http://www.wmo.ch/web/www/TEM/gts.html or by ftp from PMEL see http://www.pmel.noaa.gov/tao/proj_over/availability.html for availability http://www.pmel.noaa.gov/tao/data_deliv/deliv-pir.html
Data agreement	None
Data products provided to GHRSST-PP	Wind speed, wind direction, Air temperature, relative humidity, rainfall, http://www.pmel.noaa.gov/pirata/pir_statgus.html for full details of sensors and available data.
Data format document	http://www.pmel.noaa.gov/tao/data_deliv/deliv-pir.html
Primary GHRSST-PP entry point	TBD
Secondary GHRSST-PP entry point	
Notes	

A3.4.3 ARGO profiling floats GHRSST-PP data streams.

GHRSST-PP ID	**ARGO**
Point of contact	Sylvie Pouliquen
Data provider	1. **CORIOLIS** data centre, IFREMER, Brest France. (http://www.coriolis.eu.org/coriolis/cdc/) 2. **US-GODAE** server (http://www,usgodae.org/argo/argo.html)
Data transport	ftp pull
Data agreement	TBD

Data products provided to GHRSST-PP	**Vertical profiles of SSTdepth (below 5m), pressure and conductivity** Argo is a global array of 3,000 free-drifting profiling floats that will measure the temperature and salinity of the upper 2000 m of the ocean. This will allow continuous monitoring of the climate state of the ocean, with all data being relayed and made publicly available within hours after collection. Argo deployments began in the year 2000. The Argo array is part of the Global Climate Observing System/Global Ocean Observing System (GCOS/GOOS) and part of the Climate Variability and Predictability Experiment (CLIVAR) and the Global Ocean Data Assimilation Experiment (GODAE). GHRSST-PP will use ARGO data within the top 10m of the sea surface contributing to the MDB and validation systems
Data format document	Argo data users manual (available at http://www.coriolis.eu.org/cdc/argo_rfc.htm)
Primary GHRSST-PP entry point	RDAC/GDAC
Secondary GHRSST-PP entry point	
Notes	

A3.4.4 Radiometer GHRSST-PP data streams.

GHRSST-PP ID	**RADIOM:**
Point of contact	Peter Minnett, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway,Miami, FL 33149-1098, USA (email: pminnett@rsmas.miami.edu, Tel: Tel: +1 (305) 361-4104)
Data provider	Peter Minnett (pminnett@rsmas.miami.edu) • M-AERI measurements Craig Donlon (craig.donlon@jrc.it) • ISAR radiometer data in E-channel/Bay of Biscay Ian Robinson (ian.s.robinson@soc.soton.ac.uk) • ISAR radiometer data in E-channel/Bay of Biscay Andrew Jessup (jessup@apl.washington.edu) • CIRIMS radiometer data Ian Barton (ian.barton@marine.csiro.au) • DAR-011 and TASCO measurements Tim Nightingale (tim.nightingale@rl.ac.uk) • SISTeR measurements
Data transport	ftp pull by arrangement with individual PIs.
Data agreement	Several informal and formal agreements may be required between GHRSST-PP and the individual PI's responsible for the collection and processing of in situ radiometer data sets.
Data products provided to GHRSST-PP	SSTskin and other parameters
Data format document	Various data formats associated with different radiometer data sets. Ideally the following GHRSST-PP radiometer ASCII CSV data format is proposed: yyyy,mm,dd,hh,mm,ss,latitude,longitude,SSTskin,SSTdepth,wind_speed, solar_radiation [optional other data]
Primary GHRSST-PP	RDAC

entry point	
Secondary GHRSST-PP entry point	
Notes	Unprocessed radiometer data may be difficult to use. A real time data service for certain deployments in place (Caribbean, soon in Bay of Biscay/E. Channel) and other NRT services are in preparation (trans-Atlantic)

A3.4.5 CORIOLIS GHRSST-PP data streams.

GHRSST-PP ID	**CORIOLIS-DRIFTER** **CORIOLIS-SHIP** **CORIOLIS-PROFILE**
Point of contact	Sylvie Pouliquen
Data provider	CORIOLIS data centre, IFREMER, Brest France. (http://www.coriolis.eu.org/coriolis/cdc/)
Data transport	ftp pull from ftp://ftp.coriolis.eu.org
Data agreement	None
Data products provided to GHRSST-PP	GTS underway and profile data, buoy observations.
Data format document	http://www.coriolis.eu.org/coriolis/cdc/
Primary GHRSST-PP entry point	EURDAC
Secondary GHRSST-PP entry point	
Notes	

A3.4.6 MEDS GHRSST-PP data streams.

GHRSST-PP ID	**MEDS-DRIFTER** **MEDS-SHIP** **MEDS-PROFILE**
Point of contact	Estelle Couture (couture@meds-sdmm.dfo-mpo.gc.ca, Tel: +613 990-0259
Data provider	Marine Environmental Data Service, Department of Fisheries and Oceans Canada, 12W082 - 200 Kent Street, Ottawa, Ontario, K1A 0E6, Canada.
Data transport	GTS
Data agreement	TBD
Data products provided to GHRSST-PP	**SST1m, sea level pressure, surface current velocity** Global sea surface meteorological and oceanographic observations reported in daily and historical time frames. Drifting buoy data includes the buoy position, date, time and in most cases includes many variables such as surface and subsurface water temperature, air pressure, air temperature, wind speed and wind direction.
Data format document	MEDS is the world data centre for drifting buoys (Responsible National Oceanographic Data Centre - RNODC). As part of its role, MEDS acquires, processes, quality controls and archives real-time drifting buoy messages reporting over the Global Telecommunications System (GTS) as well as delayed mode data acquired from other sources. The real time data available from MEDS that include SST measurements are:

	SST: -Drifting buoys http://www.meds-sdmm.dfo-mpo.gc.ca/alphapro/rnodc/main_glob_e.shtml -Thermosalinograph on ships of opportunity http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Databases/TrackOB/Trackob_e.htm Temperature profiles: -BATHYs and TESACs http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Databases/OCEAN/Realtime_e.htm Data will be delivered several times per week via FTP and in the format of your choice, either ASCII or Net CDF.
Primary GHRSST-PP entry point	RDAC
Secondary GHRSST-PP entry point	TBD
Notes	Some duplication may occur between CORIOLIS and MEDS

A3.4.7 FNMOC GHRSST-PP data streams.

GHRSST-PP ID	**FNMOC-DRIFTER** **FNMOC-SHIP** **FNMOC-PROFILE**
Point of contact	Jim Cummings
Data provider	Fleet Numerical Meteorology and Oceanography Center (FNMOC)
Data transport	ftp pull from http://www.usgodae.fnmoc.navy.mil
Data agreement	GODAE
Data products provided to GHRSST-PP	Observations from ship engine room intake, fixed buoy, drifting buoy, ship hull sensors, and CMAN stations are included in the dataset.
Data format document	http://www.usgodae.fnmoc.navy.mil
Primary GHRSST-PP entry point	http://www.usgodae.fnmoc.navy.mil
Secondary GHRSST-PP entry point	
Notes	Some duplication may occur between CORIOLIS and MEDS

A3.5 GDS Numerical Weather Prediction data streams

Table 3.5.1 provides a summary of the in situ data that are available to the GHRSST-PP in near real time (2-3 day delay).

Table 3.5.1 Summary of Numerical Weather Prediction (NWP) data considered by the GDS.

GHRSST-PP Identification Code	Model	Variables	Data provider	Data Format	GHRSST-PP ingestion	Data Transport	Applicable Data Agreement
NCEP		TBD	TBD	TBD	TBD	TBD	
MetOffice		Net short-wave radiation, net long-wave radiation, latent heat flux, sensible heat flux, Precipitation evaporation Air temperature (K), SST (K), sea ice concentration (%), Air temperature (K), dew point temperature (K) Wind speed components (ms-1),	MetOffice, UK	netCDF via Grid Data Access Server (GADS) server http://www.nerc-essc.ac.uk/godiva	EURDAC/RDAC/GDAC	Data grid	Yes, MERSEA/GODAE through the GHRSST-PO. Password required.
ECMWF		SSI, wind speed, sea ice, surface thermal radiation	European Centre for Medium Range Weather Forecasting	netCDF via Grid Data Access Server (GADS) server http://www.nerc-essc.ac.uk/godiva	EURDAC/RDAC/GDAC	Data grid	Yes, GODAE, through the GHRSST-PO. Password required.
NAAPS	Navy Aerosol Analysis and Prediction System (NAAPS)	Tropospheric aerosol	Naval Research Laboratory Marine Meteorology division	Various at http://www.nrlmry.navy.mil/aerosol/	GDAC and RDAC	LAS server	None

A3.5.1 NCEP NWP GHRSST-PP data streams.

GHRSST-PP ID	NCEP
Point of contact	TBD
Data provider	TBD
Data transport	TBD
Data agreement	TBD
Data products provided to GHRSST-PP	TBD
Data format document	TBD
Primary GHRSST-PP entry point	TBD
Secondary GHRSST-PP entry point	TBD
Notes	TBD

A3.5.2 MetOffice NWP GHRSST-PP data streams.

<table>
<tr><th>GHRSST-PP ID</th><th>MetOffice</th></tr>
<tr><td>Point of contact</td><td>Dave Storkey (Dave.Storkey@metoffice.com), Tel: +44 (0)1392 884085</td></tr>
<tr><td>Data provider</td><td>Met Office, United Kingdom</td></tr>
<tr><td>Data transport</td><td>ESSC website at Reading University, UK. http://www.nerc-essc.ac.uk/godiva</td></tr>
<tr><td>Data agreement</td><td>Through GODAE/MERSEA contact Chunlei Liu at cll@mail.nerc-essc.ac.uk</td></tr>
<tr><td>Data products provided to GHRSST-PP</td><td>Surface fluxes and Near-surface parameters used to drive GODAE ocean and sea-ice models.

The following fields are currently available on the ESSC website, under “Met Office FOAM Global model”. Analyses are available under the heading “Analysis level 1” and forecasts out to T+144 hours under the heading “Forecast”. Analyses are archived back 1st Feb 2004, but forecasts are only available for today.

(See text below for elucidation of the X, L, I symbols).

<table>
<tr><td>Surface wind stresses (Nm-2)</td><td>Northward and eastward stress X components</td></tr>
<tr><td>Surface heat fluxes
(W m-2)</td><td>Net short-wave radiation L, net long-wave radiation L, latent heat flux L, sensible heat flux L, sublimation rate I surface melting flux I, conductive heat flux through bottom of sea ice I</td></tr>
<tr><td>Surface mass fluxes
(kg m-2 s-1)</td><td>Precipitation L, evaporation L, snowfall rate X</td></tr>
</table>

The following fields appear on the website under the heading “Analysis”. They are not currently available. They may be provided in the future depending on user requirements.

<table>
<tr><td>Other surface parameters</td><td>Air temperature (K), SST (K), mean sea level pressure (Nm-2), sea ice concentration (%), atmospheric stability</td></tr>
<tr><td>Parameters at 2m</td><td>Air temperature (K), dew point</td></tr>
</table>
</td></tr>
</table>

	<table><tr><td>height</td><td>temperature (K)</td></tr><tr><td>Parameters at 10m height</td><td>Wind speed components (ms-1), air temperature (K), specific humidity</td></tr></table> Some NWP models use tiling scheme in which a fraction of the surface is treated as open water and the rest is treated as sea-ice (or conceivably land). The superscripts in the table above indicate which of the surface fluxes apply to the whole of a grid box (X), only the open water (L for leads), and only the sea-ice fraction (I).
Data format document	The data is available either as GIF plots, or in netCDF format. Registration is required for netCDF downloads.
Primary GHRSST-PP entry point	EU RDAC
Secondary GHRSST-PP entry point	
Notes	Data access is available now (from 1st Feb 2004)

A3.5.3 ECMWF NWP GHRSST-PP data streams.

GHRSST-PP ID	**ECMWF**
Point of contact	Keith Haines c/o Chunlei Liu at cll@mail.nerc-essc.ac.uk
Data provider	European Centre for Medium Range Weather Forecasting (ECMWF)
Data transport	LAS/ftp ECSS web site at Reading University, UK. http://www.nerc-essc.ac.uk/godiva
Data agreement	Data agreement with GODAE, use only allowed within the GHRSST-PP. Contact Chunlei Liu at cll@mail.nerc-essc.ac.uk
Data products provided to GHRSST-PP	Global coverage 0.5° x 0.5° Forecast fields at 00:00, 06:00, 12:00, 18:00 for day D and for D+1. octet 9: parameters: 31: Sea ice cover (0-1) 165: 10 metre U wind component 166: 10 metre V wind component 176: Surface solar radiation 177: Surface thermal radiation
Data format document	Various output formats can be configured via ECSS Live access server.
Primary GHRSST-PP entry point	EU RDAC
Secondary GHRSST-PP entry point	Other RDAC/GDAC
Notes	Only available for use within the GHRSST-PP under GODAE data agreement coordinated via GHRSST-PP Project Office

A3.5.4 Navy Aerosol Analysis and Prediction System (NAAPS) GHRSST-PP data streams.

GHRSST-PP ID	**NAAPS**
Point of contact	Jim Cummings (cummings@nrlmry.navy.gov)

Data provider	Naval Research Laboratory, Marine Meteorology division
Data transport	ftp pull from http://nrlmry.navy.gov/aerosol
Data agreement	None required
Data products provided to GHRSST-PP	Global coverage as gif images
Data format document	TBD
Primary GHRSST-PP entry point	GDAC and RDAC
Secondary GHRSST-PP entry point	
Notes	Access to binary data TBD

Appendix A4 GDS match-up database (MDB)

The GHRSST-PP Matchup Database (MDB) is a collection of near contemporaneous in situ SST and GDS L2P SST data products. GHRSST-PP MDB records will be ingested into a relational database system (e.g., MySQL, ORACLE) where they can be analysed to generate Single Sensor Error Statistics (SSES). The use of a relational database means that data can be managed most effectively and made available as an on-line resource for the GHRSST-PP community to work with collectively.

The rationale underpinning the creations, population and maintenance of a common shared GHRSST-PP MDB is

1. The GHRSST-PP needs to assign an error estimate to each SST value reported in each L2P data set. The reference data set for error estimation is the in situ SST1m reported by moored and drifting buoys, ARGO, ships, research cruises and other quality controlled in situ SST data. The GHRSST-PP MDB is not restricted in terms of the SST data that are included (i.e., SST5m and SSTskin data can also be included in the database). GHRSST-PP assumes that SST data producers have generated the highest quality SST data sets possible in NRT. The GHRSST-PP MDB provides the resource necessary to derive Single Sensor Error Statistics (SSES) for a given time period, satellite sensor and SST data product; SSES are quantitative values of SST bias and standard deviation that can be assigned to L2P **`proximity_confidence`** values generated by the NRT GHRSST-PP data processors.
2. The GHRSST-PP MDB will use the Extensible Mark-up Language (XML) as a common interface for data ingestion. XML provides a convenient and extensible receptacle to transport GHRSST-PP MDB data records (matchup data and metadata) between institutions and the database itself in a common format that is easily read in ASCII format.
3. There are many MDB systems in existence that are maintained by individuals or institutions that use different time and space collocation criteria to match satellite and in situ data. Data from these systems cannot be easily compared or merged together to provide a single resource. The GHRSST-PP MDB provides a common shared resource will alleviate these problems. The time and space matchup criteria are deliberately broad in specification (within 25km and within 6 hours, based on the criteria used by current operational systems) in order to capture a large volume of data. Data analysis routines can further limit the time space criteria as required within the database itself.
4. Metadata describing the collocation data sets is often spare or non-existent making it difficult to control the quality of matchup data. The GHRSST-PP MDB design includes comprehensive metadata specifications that will greatly facilitate the quality control of MDB records.
5. Conventional MDB are stored as a data file rather than within a relational database requiring different read/write routines for each database file. The GHRSST-PP MDB will use a relational database system to store and manipulate MDB records. Data can be output in any format and all data analysis can be performed on-line within the database system itself.
6. MDB records should be submitted to the MDB system in NRT if possible although delayed mode data are also acceptable as the MDB will also be used within the GHRSST-PP reanalysis project.

The main users of the GHRSST-PP MDB will be the GHRSST-PP Science Team and the Project Office who are responsible for the regular production of SSES for the NRT operational SST data products and for the off-line reanalysis data products. **The primary output of the GHRSST-PP MDB will be SSES for each L2P data set used within the GHRSST-PP operational and reanalysis systems.** However, the MDB is not a restricted resource and will be implemented so that any registered user can access the database for their own research and for verification of the GHRSST-PP methods. **It is foreseen that other air-sea interaction projects will make use of the GHRSST-PP database system for alternative work.** For example, ARGO data within the MDB will be used to validate GDS analysis data products.

Each RDAC and GDAC centre will be responsible for preparing GDS MDB records and submitting these to the GHRSST-PP MDB system. For a given L2P satellite SST data product, in situ observations are matched at RDAC to satellite retrievals and a satellite – in situ data matchup-pair is generated. The MDB will also include in situ observations matched to L2P data records. In situ observations from 100m beneath and above the ocean surface are valid so that the structure of the atmosphere and ocean may be captured at the time of matchup. It is permitted to accept the mean average of in situ observations; for example a temporal average at a given location (mooring), the average over a vertical layer using moorings/profiler observations (e.g., for comparison to SSTfnd data products) o horizontal averages along a horizontal transect (e.g., 1s TSG observations reduced to 2 minute mean quantities). The following sections describe the content, format and submission details for GHRSST-PP MDB records.

MDB record preparation

GHRSST-PP MDB records are prepared as XML data records that are collected together in an XML MDB file. Each XML file is then sent to the GHRSST-PP GDAC where it will be checked for format consistency and ingested into the GHRSST-PP MDB system. It is not necessary to send individual XML records to the GDAC.

The basic structure of a GDS MDB data file is summarised in Table A4.1. An MDB file contains a collection of MDB records. Each MDB record has **three distinct sections**;

1. A **metadata** section providing information on the MDB record itself
2. **in situ observations** and
3. **L2P data** that is matched in space and time to the in situ observations.

An MDB data file is formatted as an XML document according to a strict Document type definition (DTD) provided below.

Table A4.1 Schematic representation of GDS MDB components

MDB file section	Description	
MDB_FILE	MDB file header (opening XML tag <MDB_FILE>)	
MDB_RECORD	**Metadata**	Data describing the data centre, contact person, release version etc.
	In situ data	Data values from in situ data collected from a single platform.
	L2P data	L2P data extracted for a m x n array centred on the location of the in situ data platform
MDB_RECORD	**additional MDB records**	
MDB_FILE	MDB file footer (closing XML tag </MDB_FILE>)	

The following sections provide a complete technical reference for the content and format of GHRSST-PP MDB data records.

A4.1 GDS rules for the generation of MDB validation records

The following GDS rules apply to the preparation of GHRSST-PP MDB records:

- **Rule 4.1.1:** A valid match-up data record is one that includes a valid satellite SST retrieval and one valid in situ observation obtained within 6 hours of each other and separated by no more than 25 km. This is called a **primary match-up**.
- **Rule 4.1.2:** The time difference between in situ observation and satellite observation should be no greater than 6[17] hours. Ideally there should be no time difference and data providers are encouraged to select the matchup data that is closest in time and space to the in situ observation.
- **Rule 4.1.3:** In the case on an analysis SST product (L4), a match-up data record includes an analysis grid point value and one in situ observation obtained within the period represented by the analysis. This is called a **primary L4 match-up**.
- **Rule 4.1.4**: Any or all satellite SST retrievals (any part or all of the pixel) located within a 25km radius **centred on the in primary match-up in situ location** and within 6 hours of the in situ observation may be included in the MDB record. This can be achieved by extracting a box around this location or selecting valid SST's retrieval within this radius. Absolute or relative position of the pixels must be provided in order to maintain the proximity pattern information.
- **Rule 4.1.5:** In the case of coarse grid (>25km) satellite data, a 3x3 array of satellite retrievals **centred on the in primary match-up in situ location** must be included in the MDB record.
- **Rule 4.1.6:** In situ data may be included from instrumentation mounted on moored buoys or profiling floats etc. up to a depth of -100m and a height of +100m (e.g., air temperature sensor, low level aircraft SSTskin profile).
- **Rule 4.1.7:** missing or invalid values should not be included within a MDB record
- **Rule 4.1.8:** If multiple in situ observations are available (e.g., from a moving ship, drifting buoy, vertical mooring) in situ observations may be averaged over a user defined window or used to define individual satellite and in situ matchup pairs (The latter case allows a more refined analysis of the MDB).
- **Rule 4.1.9:** GHRSST-PP MDB data records will be prepared and submitted as XML data files (i.e., an MDB data file contains multiple MDB records) according to the format described in Table A4.1.1 using the XML DTD provided in Table A4.3.1 and example provided in Figure A4.3.2. This format is extremely flexible allowing many diverse in situ measurements to be included in the data record as parameters that provide the data value and ancillary information to interpret the data values. Valid parameters are defined in Table A4.1.2.
- **Rule 4.1.10:** Multiple MBD records from L2P and L4 products may be gathered together into a single MDB file.
- **Rule 4.1.11:** MDB data records shall be submitted to the GHRSST-PP MDB system as XML encoded files formatted according to the GHRSST-PP XML Document Type Definition (DTD) specified in the following sections.
- **Rule 4.1.12:** Reference Tables A4.1.2, A4.1.3, and A4.1.4 provide keyword specifications for GHRSST-PP MDB data records. The GHRSST-PO will be responsible for maintaining the list of valid parameter descriptions.
- **Rule 4.1.13:** Where additional keywords are required but not listed in Reference Tables A4.1.2, A4.1.3, and A4.1.4 RDAC are encouraged to include these data as additional parameters. In this case, parameter details **must** be submitted to the GHRSST-PO and included in Table A4.1.2 for future reference (primarily to update database tables and record parse routines so that we can use the additional information).
- **Rule 4.1.14:** MDB records may be submitted to the GHRSST-PP MDB system at any time although the emphasis should be to provide suitable data within a few days of a given satellite acquisition. This will enable more accurate and regular SSES to be issued by the GHRSST-PO.
- **Rule 4.1.15:** MDB tags and attribute identifiers are all written as lowercase.

Table A4.1.2 Keyword definitions for MDB validation record as of 2005-01-11. Additional keyword definitions should be provided to the GHRSST-PO by RDAC and GDAC as required.

ID	Name	Unit of Measurement	Description	Notes
1	**sst_skin**	K	Sea surface temperature	As measured by an IR or MW radiometer
2	**sst_depth**	K	Sea temperature at depth	As measured by a contact thermometer
3	**sst_foundation**	K	Foundation temperature	Derived from an analysis system
4	**wind_speed**	m/s	Wind speed	m/s
5	**wind_direction**	Degrees true	Wind direction	Degrees
6	**air_temp**	K	Air temperature	K
7	**rel_humidity**	%	Relative Humidity	%
8	**twet_bulb**	K	Wet bulb air temperature	K
9	**tdry_bulb**	K	Dry bulb air temperature	K
10	**qsolar**	W/m^2	Solar radiation	W/m2
11	**qlw_down**	W/m^2	Downwelling longwave radiation	W/m2
12	**qlatent**	W/m^2	Latent heat flux	W/m2
13	**qsensible**	W/m^2	Sensible heat flux	W/m2
14	**qnet_longwave**	W/m^2	Longwave net heat flux	W/m2
15	**surf_current**	m/s	Surface current velocity	m/s
16	**surf_current_dir**	Degrees true	Surface current direction	Degrees true
17	**precip**	Mm/hr	Precipitation rate	mm/hr
18	**cloud_cover**	Oktas	Cloud cover	Oktas
19	**sig_wave_ht**	M	Significant wave height	M
20	**swell_dir**	Degrees true	Swell direction	Degrees true
21	**salinity**	psu	Salinity	
22	**par**	W/m^2	Photosynthetically Active Radiation	
23	**platform_SMG**	knots	speed made good of measurement platform	
24	**platform_CMG**	degrees true	course made good of measurement platform	
24	**ver_loc_measurement**	metres	Vertical location of measurement; +tive above ocean surface .e., height of air_temp measurement, depth of SST measurement	
25	**RDAC defined – please inform GHRSST-PO**			

Table A4.1.3 Definition MDB platform keywords as of 2005-01-11. Additional keyword definitions should be provided to the GHRSST-PO by RDAC and GDAC as required.

ID	Keyword	Description
1	**research_ship**	Water going vessel
2	**moored_buoy**	
3	**drifting_buoy**	
4	**aircraft**	
5	**balloon**	
6	**glider**	
7	**shore_station**	
8	**light_vessel**	
9	**vos**	
10	**polar_satellite**	e.g., AVHRR
11	**geostationary_satellite**	e.g., MSG or GOES
12	**low_earth_orbit_satellite**	e.g., TRMM
13	**float**	
…	**RDAC defined – please inform GHRSST-PO**	

Table A4.1.4 Definition MDB units keywords as of 2004-01-11. Additional keyword definitions should be provided to the GHRSST-PO by RDAC and GDAC as required.

ID	Keyword	Description
1	**Kelvin**	Kelvin temperature scale
2	**celsius**	Celsius temperature scale
3	**watt_m-2**	Watts per meter squared

4	**m_s-2**	Metres per second
5	**percent**	Percentage
6	**mm_hr-2**	Millimetres per hour
9	**Ok**	Okta
10	**deg_t**	Angular degrees true
11	**Deg_r**	Angular degrees relative
12	**RDAC defined – please inform GHRSST-PO**	

A4.2 MDB file naming convention

An MDB filename shall comply to the following file naming convention:

`MDB-<Creation Date>-<Satellite file name>.<base format>`

Table A4.2.1 Filename convention components for GHRSST-PP MDB data files.

Name	Definition	Description
<Creation Date>	yyyymmddThhmmssZ	The creation date of the MDB records contained in this file (several MDB files may be created successively for the same satellite file since some in situ observations will be delivered lately). Time specified in UTC
<satellite filename>	Related satellite filename (without the format extension)	The name of the data file from which the satellite pixels of the MDB records are extracted.
<base format>	xml	Generic file format

For example

`MDB-20040227T123458Z-EUR-L4UHfnd-MED-v01.xml`

Would refer to a MDB data file created on the 22nd February 2004 at 12:34:58 UTC and contains in situ data matched to EUR-SSTUHfnd-MED data produced by the GDS v1.0.

4.3 MDB XML document type definition (DTD)

GHRSST-PP MDB records shall be formatted in XML format, complying to the following DTD (note that all tags and qttributes are lowercase and are case sensitive). The mdb.dtd can be found at http://ghrsst-pp.metoffice.gom/mdb.dtd

```
<!-- DTD for the GHRSST-PP MDB                    -->
<!--                                                    -->
<!-- Author: J. F Piolle and C J Donlon                -->
<!--                                                    -->
<!-- History                                            -->
<!-- 2004-02-17                                 -->
<!-- 2005-01-11                                 -->

<!ELEMENT mdb_file ( mdb_record+ )>

<!ELEMENT mdb_record ( version,
            data_center,
            creation_date,
            last_revision_date?,
            revision_history?,
            acquisition,
            acquisition
          )
```

```
>

<!ELEMENT version (#PCDATA )>


<!-- Data Center ================================================== -->
<!ELEMENT data_center ( data_center_name,
                    data_center_url*,
                    data_set_id?,
                    personnel+
                  )
>

<!-- Data Center Name ============================================= -->
<!ELEMENT data_center_name ( short_name,
                         long_name
                        )
>
<!ELEMENT short_name (#PCDATA)>
<!ELEMENT long_name (#PCDATA)>

<!ELEMENT data_center_url (#PCDATA)>
<!ELEMENT data_set_id (#PCDATA)>

<!-- Personnel ==================================================== -->
<!ELEMENT personnel ( role+,
                  first_name,
                  middle_name?,
                  last_name,
                  email,
                  phone,
                  fax?,
                  address
                )
>

<!ELEMENT role (#PCDATA)>
<!ELEMENT first_name (#PCDATA)>
<!ELEMENT middle_name (#PCDATA)>
<!ELEMENT last_name (#PCDATA)>
<!ELEMENT email (#PCDATA)>
<!ELEMENT phone (#PCDATA)>
<!ELEMENT fax (#PCDATA)>
<!ELEMENT address (#PCDATA)>


<!ELEMENT creation_date (#PCDATA )>
<!ELEMENT last_revision_date (#PCDATA )>
<!ELEMENT revision_history (#PCDATA )>

<!-- acquisition ================================== -->

<!ELEMENT acquisition ( dataset,
                    platform,
                    parameter+
                  )
>

<!ATTLIST acquisition type ( in_situ |
                         satellite
                        ) #REQUIRED
>


<!ELEMENT dataset  (#PCDATA)>

<!-- platform ========================================== -->

<!ELEMENT platform ( identifier?,
                 context*)>

<!ATTLIST platform type ( ship |
                      drifting_buoy |
                      moored_buoy |
                      aircraft |
```

```
                    balloon |
                    glider |
                    shore_station |
                    light_vessel |
                    polar_satellite |
                    geostationary_satellite |
                    float
                   ) #REQUIRED
>

<!ELEMENT identifier  (#PCDATA)>

<!-- context =========================================== -->

<!ELEMENT context (#PCDATA)>

<!ATTLIST  context parameter ( heading |
                        speed |
                        course |
                        satellite_zenith_angle |
                        dkm_cloud |
                        dkm_land |
                        dkm_rain |
                        dkm_ice
                       ) #REQUIRED
>

<!ATTLIST  context unit ( angular_degree |
                    knot |
                    degree_true
                   ) #REQUIRED
>

<!-- parameter =========================================== -->

<!ELEMENT parameter (observation+)>


<!ATTLIST parameter type ( sea_surface_temperature |
                     sst_dtime |
                     sses_bias_error |
                     sses_standard_deviation_error |
                     dt_min |
                     surface_solar_irradiance |
                     sea_ice_fraction |
                     aerosol_optical_depth |
                     proximity_confidence |
                     air_temperature |
                     wind_speed |
                     wind_direction |
                     dew_point_temperature
                    ) #REQUIRED
>

<!ATTLIST parameter unit ( kelvin |
                     m_s-1 |
                     celsius |
                     watt_m-2 |
                     percent |
                     ghrsst_confidence_table |
                     angular_degree |
                     um
                    ) #REQUIRED
>

<!ATTLIST parameter sensor ( avhrr |
                      aatsr |
                      amsr-e |
                      seviri |
                      tsg
                     ) #IMPLIED
>

<!ATTLIST parameter algorithm CDATA  #IMPLIED>
<!ATTLIST parameter space_resolution CDATA  #IMPLIED>
```

```
<!ATTLIST parameter box (1x1 | 3x3 | 5x5 | 15x15 | 25x25) #IMPLIED>
<!ATTLIST parameter radius CDATA  #IMPLIED>


<!-- observation========================================== -->

<!ELEMENT observation ( latitude,
                        ,
                        _abs_x?,
                        _abs_y?,
                        time,
                        _location?,
                        _sample_interval?,
                        _standard_deviation?,
                        _start_lat?,
                        _start_lon?,
                        _start_time?,
                        _stop_lat?,
                        _stop_lon?,
                        _stop_time?,
                        ,
                        _bias_error?,
                        _standard_deviation_error?)>

<!ATTLIST observation type ( single |
                      point_average |
                      line_average |
                      vertical_average
                     ) #REQUIRED
>
<!ATTLIST observation pixel( validation|
                      context
                     ) #IMPLIED
>


<!ELEMENT latitude  (#PCDATA)>
<!ELEMENT longitude  (#PCDATA)>
<!ELEMENT box_abs_x (#PCDATA)>
<!ELEMENT box_abs_y (#PCDATA)>
<!ELEMENT date-time  (#PCDATA)>
<!ELEMENT vertical_location (#PCDATA)>

<!ELEMENT average_sample_interval (#PCDATA)>
<!ATTLIST average_sample_interval sample_unit CDATA #REQUIRED>

<!ELEMENT average_standard_deviation (#PCDATA)>
<!ELEMENT average_start_lat (#PCDATA)>
<!ELEMENT average_start_lon (#PCDATA)>
<!ELEMENT average_start_time (#PCDATA)>
<!ELEMENT average_stop_lat (#PCDATA)>
<!ELEMENT average_stop_lon (#PCDATA)>
<!ELEMENT average_stop_time (#PCDATA)>
<!ELEMENT value (#PCDATA)>
<!ELEMENT value_bias_error (#PCDATA)>
<!ELEMENT value_standard_deviation_error (#PCDATA)>
```

A4.3.1 MDB XML DTD description (2005-01-11, v2.3)

Table A4.3.1 provides a more explicit and readable description of the GHRSST-PP MDB DTD and is fully consistent with the DTD described above. It provides a full description of each field and associated attributes that should be used to develop an XML format MDB data file. Some users will prefer to use this table as reference instead of the DTD. An example MDB data file is given in the following section.

Field Name and attribute description	Cardinality	Type	Description
mdb_file	1	group	'mdb-file' is the envelop of a set of MDB records.
mdb_record	1+	group	'mdb_record' is the envelop of the record (several records may be contained within a file)
data_center	1	group	This group identifies the data centre which distributes the data, the data centre URL, data set identification, and a person to contact. Note USA spelling of centre
data_center_name	1	group	Name of the data centre taken from the list of valid keywords http://gcmd.gsfc.nasa.gov/Resources/valids/data_centre.html
short_name	1	Keyword	Name of the data centre taken from the list of valid keywords http://gcmd.gsfc.nasa.gov/Resources/valids/data_centre.html
long_name	1	Keyword	Name of the data centre taken from the list of valid keywords http://gcmd.gsfc.nasa.gov/Resources/valids/data_centre.html
data_center_url	0+	string	The Internet Uniform Resource Locator(s) (URL) for the data centre should be listed if available. This field may be repeated as many times as necessary within the data centre group. The URLs will be hypertext linked. The URLs may include gopher, ftp and telnet as well as WWW servers e.g., Data_Center_URL: http://daac.gsfc.nasa.gov Data_Center_URL: ftp://podaac.jpl.nasa.gov Data_Center_URL: telnet://ncar.ucar.edu Data_Center_URL: gopher://www.ciesin.org
data_set_id	0|1	string	Should be listed if available. These are identification codes assigned by the data centre, which will simplify the location and ordering of the data sets.
personnel	1+	group	This group describes the primary contact person at the data centre who is able to respond to requests for and queries about the data. Most fields are self-explanatory. Address should be a valid postal address for the named person. Email, Phone and FAX fields are optional and may be repeated. These contact personnel include Investigator, Technical Contact, or DSD Author.

	role	1+	Keyword	The Role may be repeated per personnel group. In other words, a person may have more than one role. In this case, the Personnel group needs only to be specified once in the DSD with repeating Role fields. Investigator: Person who headed the investigation or experiment that resulted in the acquisition of the data described (i.e., Principal Investigator, Experiment Team Leader) Technical Contact: A person who is knowledgeable about the technical content of the data (quality, processing methods, units, available software for further processing) DSD Author: The person who is responsible for the content of the DSD. If the responsibility shifts from the original author to another person, the DSD Author field should be updated to the new responsible person
	first_name	1	string	Primary contact name
	middle_name	0 | 1	string	Primary contact name
	last_name	1	string	Primary contact name
	email	1	string	Primary contact name
	phone	1	string	Phone and fax numbers should be easily read by users. For example, for USA phone numbers, use either (xxx) xxx-xxx or xxx-xxx-xxxx. For other countries, begin the number with the country code: +44 123 45678
	fax	0 | 1	string	
	address	1	string	Primary contact address. The country should always be included in the address.
	version	1	int	The version of this metadata record. Version 1 is the original and File_Version number should be incremented for each new release of the metadata.
	creation_date	1	string	The date the metadata was created. Years should be four-digits. Month and day of month should be two digits, with leading zeroes if necessary
	last_revision_date	0 | 1	string	The date the metadata was last revised. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC

	revision_history	0|1	string	A listing of changes made to over time. Provides a mechanism for tracking revisions to DSD content.
	acquisition type | Req | string | in_situ satellite	2	group	Details of a data acquisition (e.g. an in situ or satellite measurement). The 'type' attribute specifies the kind of observation (in situ or satellite)
	dataset	1	string	Identifier of the dataset (use dsd_entry code for satellite)
	platform type | Req | string | ship moored_buoy drifting_buoy aircraft balloon glider shore_station light_vessel float polar_satellite geostationary_satellite	1	group	Details of measurement platform The 'type' attribute is extracted from Table A4.1.3. This list may need to be updated.
	identifier	0|1	string	Call sign or ID of platform/satellite (refer to http://gcmd.gsfc.nasa.gov/Resources/valids/sources.html for satellites) when available
	context parameter | Req | string | heading speed course satellite_zenith_angle dkm_cloud dkm_land dkm_rain dkm_ice unit | Req | string | knot degree_true angular_degree	0+	float	The context elements provide any available information related to the attitude of the measurement platform when the observation was made : this may include the speed (in knots), heading (in degree true), course (in degree true) for ships, the satellite zenith angle for a satellite. This list may be updated if additional context parameter are available. Optional parameters related to the conditions at observation points may also be inserted here such as : Dkm_cloud : Distance in Km from nearest cloud flagged data from 5x5 array centre pixel (if >100km set to 100) Dkm_land : Distance in Km from nearest land flagged data from box array centre pixel defined in the parameter section below (if >100km set to 100) Dkm_rain : Distance in Km from nearest rain flagged data from 5x5 array centre pixel (if >100km set to 100) Dkm_ice : Distance in Km fro nearest sea ice flagged data from 5x5 array centre pixel (if >100km set to 100) For unit convention, refer to GDS table A4.1.4

<table>
<tr>
<td></td>
<td>parameter
<table>
<tr><td>type</td><td>Req</td><td>string</td><td>Keyword</td></tr>
<tr><td>unit</td><td>Req</td><td>string</td><td>keyword</td></tr>
<tr><td>sensor</td><td>Opt</td><td>string</td><td>keyword</td></tr>
<tr><td>algorithm</td><td>Opt</td><td>string</td><td>value</td></tr>
<tr><td>space_resolution</td><td>Req</td><td>float</td><td>value</td></tr>
<tr><td>box</td><td>Opt</td><td>string</td><td>3x3 | 5x5 | 15x15 | 25x25</td></tr>
<tr><td>radius</td><td>Opt</td><td>float</td><td>value</td></tr>
</table>
</td>
<td>1+</td>
<td>group</td>
<td>This group contains all available valid measurements for a given variable. This may include the measurements available at various depth (for in situ measurements) or within a fixed-size box (for satellite measurements).
The 'type' attribute refers to the standard name of the variable. Refer to table A4.1.2 for in situ measurements or to the list of netCDF variable for L2P/L4 products.
For units, refer to S. I. Measurement unit table A4.1.2 and A4.1.4
'sensor' is the a n optional attribute indicating the sensor used to measure the parameter; refer to http://gcmd.gsfc.nasa.gov/Resources/valids/sensors.html for a complete list of available keywords.
'algorithm' refers to the version of algorithm
The spatial resolution is nominal size of the satellite pixels in km
'box' defines the size of the box containing the satellite pixels. GDS-1.5 specifies a 25km radius so the box size should match this requirement (25x25 for AATSR, 3x3 for TMI/AMSRE,…)
'radius' can be used instead of box. GDS-1.5 specifies a 25km radius so this should be possible and promoted as the default</td>
</tr>
<tr>
<td></td>
<td>observation
<table>
<tr><td>type</td><td>Req</td><td>string</td><td>single
point_average
line_average
vertical_average</td></tr>
<tr><td>pixel</td><td>Req</td><td>String</td><td>validation
context</td></tr>
</table>
</td>
<td>1+</td>
<td>float
int
bool</td>
<td>This group contains all information related to a valid measurement.
'type' indicates the type of averaging, if any :
'single' indicates a single value (no averaging)
'point_average' indicates that the value is an average value at a fixed point (moored buoys) e.g. temporal averaging.
'line_average' indicates that the the value is an average value along a horizontal transect [e.g., ship track] centred at [latitude, longitude,time]
'vertical_average' indicates that the the value is an average value along a vertical depth [e.g., Mooring] centred at [latitude, longitude,time,vertical_location]
'pixel' indicates if a satellite pixel is the validation pixel (the closest valid measurement to an in situ point) or a context pixel (neighbour of the validation pixel). This attribute is mandatory for satellite pixel only (not used for in situ observations).</td>
</tr>
</table>

	latitude	1	float	Latitude of observation (in situ) or pixel (satellite, reference position = centre of pixel) Positive North of equator.
	longitude	1	float	Longitude of observation (in situ) or central 5x5 array pixel longitude, (satellite, reference position = centre of pixel) Positive East of 0°.
	box_abs_x	0|1	int	Absolute x position (row number starting from 0) of the pixel in the data file. To be used when extracting a box. Can be native pixel row position. The purpose is to have a means to look up the relative location of each pixel to another.
	box_abs_y	0|1	int	Absolute y position (column number starting from 0) of the pixel in the data file. To be used when extracting a box. Can be native scan line number. The purpose is to have a means to look up the relative location of each pixel to another.
	date-time	1	string	Date and Time of acquisition. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
	vertical_location	0|1	float	vertical_location' is the height/depth of measurement specified in meters from datum at the sea surface. Negative indicates depth. Mandatory for in situ data.
	average_sample_interval sample_unit | Req | keyword | keyword	0|1	float	'sample_interval' indicates the data sample time interval of the data used to define mean_average sample_units: unit for sample_interval. For units, refer to S. I. Measurement unit table A4.1.2 and A4.1.4 For sample_unit, refer to S. I. Measurement unit table A4.1.2 and A4.1.4
	average_standard_deviation	0|1	float	'standard_deviation' is the standard deviation of an average measurement and should be used when a mean_average is provided as the value
	average_start_lat	0|1	float	'average_start|stop_lat|lon' indicates the average start position of the data used to define the value mean_average
	average_start_lon	0|1	float	

	average_start_time	0|1	string	'average_ start|stop_time' Date and Time of line transect end points. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
	average_stop_lat	0|1	float	'average_start|stop_lat|lon' indicates the average start position of the data used to define the value mean_average
	average_stop_lon	0|1	float	
	average_stop_time	0|1	string	'average_ start|stop_time' Date and Time of line transect end points. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
	value	1	float	Measurement value
	value_bias_error	0|1	float	Bias error of the value if available
	value_ standard_deviation _error	0|1	float	standard deviation error of the value if available

A4.4 Example XML formatted MDB record

The following code provides an example GHRSST-PP MDB data record formatted according to the DTD provided in A4.3.


```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE MDB SYSTEM "http://ghrsst-pp.metoffice.com/mdb.dtd">
<mdb_file>
  <mdb_record>
    <version>1.0</version>
    <data_center>
      <data_center_name>
        <short_name>IFREMER</short_name>
        <long_name>Institut Francais de Recherche pour l'Exploitation de la Mer / Centre ERS d'Archivage et de Traitement</long_name>
      </data_center_name>
      <data_center_url>http://www.ifremer.fr/cersat</data_center_url>
      <personnel>
        <role>Technical Contact</role>
        <first_name>Jean-Francois</first_name>
        <last_name>Piolle</last_name>
        <email>jfpiolle@ifremer.fr</email>
        <phone>555-333444</phone>
        <fax>555-3333222</fax>
        <address>IFREMER CERSAT, technopole Brest-Iroise, 29280 Plouzane, France</address>
      </personnel>
    </data_center>
    <creation_date>20050130T122300Z</creation_date>
    <last_revision_date>20050130T122300Z </last_revision_date>
    <revision_history>2005-01-30,First Release</revision_history>
    <acquisition type="satellite">
      <dataset>EUR-L2P-AMSRE</dataset>
      <platform type="polar_satellite">
        <identifier>AQUA</identifier>
      </platform>
      <parameter type="sea_surface_temperature" unit="kelvin" sensor="amsr-e" space_resolution="25" box="3x3">
        <observation type="single" pixel="validation">
          <latitude>36.125000</latitude>
          <longitude>-34.375000</longitude>
          <box_abs_x>287</box_abs_x>
          <box_abs_y>128</box_abs_y>
          <date-time>20050110T223544Z </date-time>
          <value>292.15</value>
        </observation>
        <observation type="single" pixel="context">
          <latitude>36.375000</latitude>
          <longitude>-34.125000</longitude>
          <box_abs_x>286</box_abs_x>
          <box_abs_y>129</box_abs_y>
          <date-time>20050110T223544Z </date-time>
          <value>292.10</value>
        </observation>
        <observation type="single" pixel="context">
          <latitude>36.375000</latitude>
          <longitude>-34.375000</longitude>
          <box_abs_x>287</box_abs_x>
          <box_abs_y>129</box_abs_y>
          <date-time>20050110T223544Z </date-time>
          <value>292.05</value>
        </observation>
      </parameter>
      <parameter type="wind_speed" unit="m_s-1" sensor="amsre-e" space_resolution="25" box="3x3">
        <observation type="single" pixel="validation">
          <latitude>36.125000</latitude>
          <longitude>-34.375000</longitude>
          <box_abs_x>287</box_abs_x>
          <box_abs_y>128</box_abs_y>
```

```
      <date-time>20050110T223544Z </date-time>
      <value>11</value>
    </observation>
    <observation type="single" pixel="context">
      <latitude>36.375000</latitude>
      <longitude>-34.125000</longitude>
      <box_abs_x>286</box_abs_x>
      <box_abs_y>129</box_abs_y>
      <date-time>20050110T223544Z </date-time>
      <value>11</value>
    </observation>
    <observation type="single" pixel="context">
      <latitude>36.375000</latitude>
      <longitude>-34.375000</longitude>
      <box_abs_x>287</box_abs_x>
      <box_abs_y>129</box_abs_y>
      <date-time>441504358</date-time>
      <value>12</value>
    </observation>
   </parameter>
  </acquisition>
  <acquisition type="in_situ">
   <dataset>CO_BA</dataset>
   <platform type="moored_buoy">
    <identifier>WAUW</identifier>
   </platform>
   <parameter type="sea_temperature" unit="kelvin">
    <observation type="single">
      <latitude>36.000000</latitude>
      <longitude>-34.400002</longitude>
      <date-time>20050110T223544Z</date-time>
      <vertical_location>-2</vertical_location>
      <value>292.05</value>
       <value_bias_error>0.1</value_bias_error>
       <value_standard_deviation_error>0.01</value_standard_deviation_error>
    </observation>
   </parameter>
  </acquisition>
 </mdb_record>
</mdb_file>
```


A4.5 MDB validation records for satellite-to-satellite comparisons

Satellite to Satellite MDB records are not required in the GDSv1.0.

Appendix A5. GHRSST-PP High Resolution Diagnostic Data Set (HR-DDS) data granule data product format specification

The GHRSST-PP High-resolution Diagnostic Data Set (HR-DDS) system provides a distributed data resource and a framework to analyse and inter-compare L2P and L4 data products in near real time, together with other data products including NWP analyses, operational ocean and atmospheric model outputs and in situ observations. In order to reduce the overhead of data storage and transport, the HR-DDS consists of about150 globally distributed 2° x 2° latitude x longitude sites. Figure A5.1 shows a map of primary[18] HR-DDS sites (v2.3) for which all L2P and L4* data streams[19] will be extracted and archived as ***data granules***. Each HR-DDS site is strategically positioned in order to address a particular issue; for example, a particularly dynamic or "quiet" oceanographic or atmospheric region, location of additional in situ infrastructure, areas known to be influenced by atmospheric aerosol or persistent cloud cover and areas already having a significant scientific interest. Each site shown in Figure A5.1 is defined in Section A5.3.

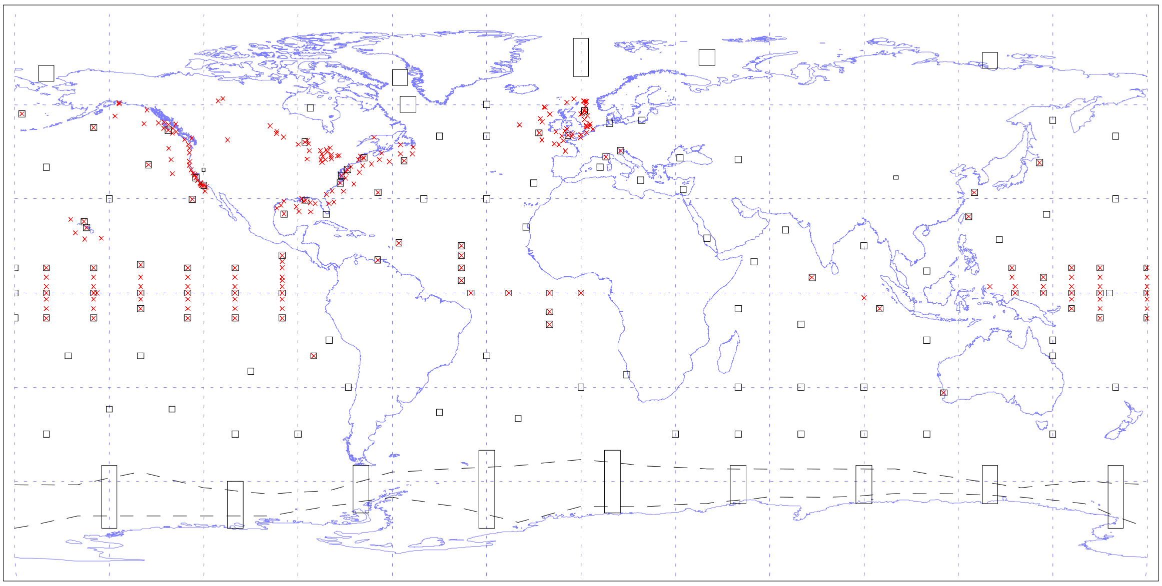


Figure A5.1 Location of HR-DDS sites v2.3 (April 2003) Based on output of the 2nd & 3rd GHRSST-PP workshop Science Team feedback. Moorings and permanent in situ installations are indicated by red cross (note that some are located on inland lakes). HR-DDS sites do not always contain in situ measurement systems. The HR-DDS is fully documented in the HR-DDS Implementation Plan (GHRSST/14).

The HR-DDS system provides a real time data resource the main learning tool for GHRSST-PP project scientists to develop and refine the GDS, to investigate differences between complementary data streams, to investigate regional and time variant bias statistics, to monitor the quality of input and output data streams and as a core component of the GHRSST-PP reanalysis (RAN) project. The HR-DDS constitutes the virtual laboratory of the GHRSST-PP allowing, for example, a full implementation of Integrated Global Observing Strategy (IGOS) measurement principles (see http://www.igospartners.org/).

This Appendix provides a summary overview of the GHRSST-PP HR-DDS system. A complete Scientific and technical description of the HR-DDS system can be found in GHRSST/13 available at http://www.ghgrsst-pp.org.

A5.1 Summary of the HR-DDS system configuration and operation

HR-DDS granules are produced by RDAC and GDAC centres in real time as netCDF data files according to the specifications provided in Appendix A1. HR-DDS granules are archived locally together with a corresponding HR-DDS metadata record. HR-DDS metadata records are produced according to the specifications described in Appendix A6 and are identical to other GDS metadata records. HR-DDS archive centres are referred to as HR-DDS nodes and in general, HR-DDS nodes should be part of RDAC and GDAC.

Figure A5.2 provides a schematic diagram of the GHRSST-PP HR-DDS data system which is based on a number of distributed HR-DDS nodes that are interconnected using the Distributed Oceanographic Data System (DODS, see http://www.dods.org).

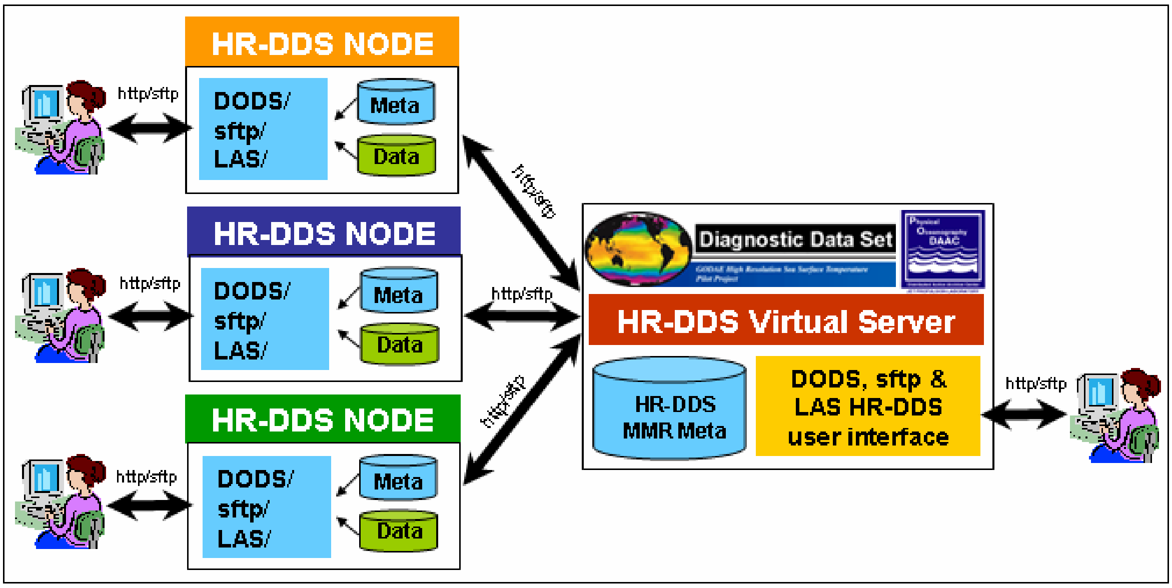


Figure A5.2 Schematic diagram of the GHRSST-PP HR-DDS system. Users may access HR-DDS data through local HR-DDS nodes or the global distributed data archive via a HR-DDS virtual server that coordinates GHRSST-PP HR-DDS data through the HR-DDS MMR.

The DODS architecture uses a client/server model and the http protocol to provide a framework that simplifies all aspects of scientific data networking. It provides tools (such as DODS servers) that make local data accessible to remote locations regardless of local storage format. HR-DDS granules at HR-DDS nodes may be accessed directly using a number of DODS aware client applications i(e.g., IDL, MATLAB) from any location. The Open source project to develop and extend a data access protocol (OPeNDAP) has now evolved from the DODS.

In addition to local data access via DODS/OPeNDAP servers at each HR-DDS node, HR-DDS nodal data archives are virtually combined as a single data archive by a HR-DDS virtual server. The HR-DDS virtual server uses GHRSST-PP Master Metadata Repository (MMR) HR-DDS metadata records to catalogue HR-DDS data granules archived at ***all*** distributed HR-DDS nodes. Using a HR-DDS interface, users at the HR-DDS virtual server may access all HR-DDS granules (maintained at distributed locations, served by local DODS servers) using DODS aware client applications as if they were a single data archive residing on a local computer.

In addition to DODS access, HR-DDS granules may also be accessed using ftp or secure ftp (sftp) transactions or through a Live Access Server (LAS, http://ferret.pmel.noaa.gov/Ferret/LAS/). LAS is a highly configurable Web server designed to provide flexible access to geo-referenced scientific data and can present distributed data sets as a unified virtual data base through the use of DODS networking Using a web browser interface LAS enables a user to

- visualize data with on-the-fly graphics

- request custom subsets of variables in a choice of file formats
- access background reference material about the data (metadata)
- compare (difference) variables from distributed locations
- LAS enables the data provider to
- unify access to multiple types of data in a single interface
- create thematic data servers from distributed data sources
- offer derived products on the fly
- remedy metadata inadequacies (poorly self-describing data)
- offer unique products (e.g. visualization styles specialized for the data)

LAS will be used extensively in the GODAE project and will allow data to be extracted from many operational ocean model systems for inclusion and comparison with HR-DDS SST data. In addition, the LAS system is in active development to provide advanced features such as temporal aggregation of data and advanced display options in collaboration with the GODAE Data Sharing Pilot Project.

Each HR-DDS node is thus responsible for:

(a) Installing and maintaining a data archive of HR-DDS granules and associated MMR DSD and MMR-FS metadata records (see Appendix A6)
(b) Operating a DODS server to serve HR-DDS data
(c) Operating an ftp/sftp server for HR-DDS granule access
(d) Operating other optional data interface software (e.g., LAS)

A5.2 HR-DDS data granule file format and filename convention

A5.2.1 HR-DDS granule filename convention

HR-DDS filenames will conform to the following format

HRDDS_<Creation Date>_<RDAC code>_<data set>_<short name>_<long name>.<format>

Table A5.2.1 Filename convention components for GHRSST-PP HR-DDS data files.

Name	Definition	Description
<Creation Date>	yyyymmddThhmmssZ yyyy=year mm=month (1-12) dd=day of month (1-31) T=time string identifier hh=hour of the day (0-23) mm=minute of the hour (0-59) ss=second of the minute (0-59) Z=indicates time specification is UTC.	The starting date/time of the data contained in this file. Time specified in UTC
<RDAC Code>	Table A2.1	The name of the RDAC/GDAC centre producing the HR-DDS data
<data set >	Table A2.2	GDS data set code name
<short name>	Table A5.4.1	Short name of the HR-DDS area covered by the data in this data file

<long name>	Table A5.4.1	Long name of the HR-DDS area covered by the data in this data file
<format>	Format of data contained within this data file	.nc for a netCDF data file and .png as a quicklook image file

e.g.,

HRDDS_20040228T002133Z_EUR_ ATS_NR__2P_SST_GHR025_GHRSST-Norfolk-Island.nc

would refer to a GDS HR-DDS data file containing data collected on the 28th February 2004 starting at 00:21:33 UTC. The HR-DDS data file was generated by the European RDAC and contains ENVISAT AATSR NR__2P 1km SST data over the HR-DDS site 'GHRSST-Norfolk-Island' and is formatted as a netCDF file.

A MMR DSD will be registered at the GHRSST-PP MMR for each HR-DDS **site** using the **HRDDS long name** that is described in Appendix A5.4 following the procedure described in Appendix A6. A MMR_FR will be generated for each HR-DDS granule and registered at the MMR following the procedure described in Appendix A6.4. Note that this arrangement is different form other data sets within the GDS where MMR_FR point to data files of the same generic type; for the HR-DDS the MMR_DSD will describe the HR-DDS site and the MMR_FR will point to many different types of GDS data products. HR-DDS filenames have been structured so that users may refer to the DSD describing the generic L2P data sets from which HR_DDS L2P data granules have been extracted and reformatted.

A5.2.2 HR-DDS granule data files

HR-DDS granules are formatted as netCDF data files according to the format description laid out in a similar style to the L2P data files described in Appendix A1. The main differences between L2P data and HR-DDS data files are:

1. HR-DDS data have been re-sampled using a nearest neighbour scheme to a uniform resolution grid specified in Table A1.1.3.
2. All L2P confidence data are re-sampled in the same manner to provide a multiple image netCDF file.
3. Coordinate variables are used to describe the uniform grid.

The HR-DDS format is defined as follows. L2P derived HR-DDS granules shall adhere to the specification of A1.2.2. All L4UHR HR-DDS granules shall adhere to the specification in A1.3.2. In both cases the attributes shall adhere to the specification in A1.1.2. In both cases the Global attributes shall adhere to the specification of A1.1.1 with the following addition.

Tag name	Format	Description
Location	string	HR-DDS short name location(eg GHR025)

All HR-DDS files shall be gridded to the specification in A1.1.3.1

A5.2.3 HR-DDS granule quicklook image files

A browse image file should also be created of the SST data only as a scaled portable network graphics (png) format file. In this case, data should be scaled to a resolution of 0.15K and an offset of -2.0 (i.e., from -2.0°C to 36.4°C). No colour tables should be applied to the data other than a 256 shade greyscale.

A5.2.4 Sample HR-DDS (CDL Header of L2P derived HR-DDS)

```
netcdf hr-dds example {
        dimensions:
                ni = 200 ;
                nj = 200 ;
                time = 1 ;

        variables:
        float lat(nj) ;
                lat:long_name = "latitude" ;
                lat:units = "degrees_north" ;
                lat:add_offset = 43.37 ;
                lat:scale_factor = 0.01 ;
        float lon(ni) ;
                lon:long_name = "longitude" ;
                lon:units = "degrees_east" ;
                lon:add_offset = 7.9 ;
                lon:scale_factor = 0.01 ;
        long time(time) ;
                time:long_name = "reference time of sst file" ;
                time:units = "seconds since 1981-01-01 00:00:00" ;
        short sea_surface_temperature(time, nj, ni) ;
                sea_surface_temperature:long_name = "sea surface temperature" ;
                sea_surface_temperature:units = "kelvin" ;
                sea_surface_temperature:_FillValue = -32768s ;
                sea_surface_temperature:add_offset = 273.15 ;
                sea_surface_temperature:scale_factor = 0.01 ;
                sea_surface_temperature:valid_min = -5000s ;
                sea_surface_temperature:valid_max = 5000s ;
                sea_surface_temperature:coordinates = "lon lat" ;
                sea_surface_temperature:grid_mapping = "cylindrical" ;
                sea_surface_temperature:source = " name of L2 source  (ex:EUMETSAT SAF O&SI)" ;
        short sst_dtime (time, nj, ni) ;
                sst_dtime:long_name = "time difference from reference time" ;
                sst_dtime:units = "second" ;
                sst_dtime:_FillValue = -32768s ;
                sst_dtime:add_offset = 0. ;
                sst_dtime:scale_factor = 1. ;
                sst_dtime:valid_min = -32767s ;
                sst_dtime:valid_max = 32767s ;
                sst_dtime:coordinates = "lon lat" ;
                sst_dtime:grid_mapping = "cylindrical" ;
        byte SSES_bias_error (time, nj, ni) ;
                SSES_bias_error:long_name = "SSES bias error based on confidence flags" ;
                SSES_bias_error:units = "kelvin" ;
                SSES_bias_error:_FillValue = -128 ;
                SSES_bias_error:add_offset = 0. ;
                SSES_bias_error:scale_factor = 0.01 ;
                SSES_bias_error:valid_min = -127 ;
                SSES_bias_error:valid_max = 127 ;
                SSES_bias_error:coordinates = "lon lat" ;
                SSES_bias_error:grid_mapping = "cylindrical" ;
        byte SSES_standard_deviation_error (time, nj, ni) ;
                SSES_standard_deviation_error:long_name = "SSES standard deviation error based on confidence flags" ;
                SSES_standard_deviation_error:units = "kelvin" ;
                SSES_standard_deviation_error:_FillValue = -128 ;
                SSES_standard_deviation_error:add_offset = 100. ;
                SSES_standard_deviation_error:scale_factor = 0.01 ;
                SSES_standard_deviation_error:valid_min = -127 ;
                SSES_standard_deviation_error:valid_max = 127 ;
                SSES_standard_deviation_error:coordinates = "lon lat" ;
                SSES_standard_deviation_error:grid_mapping = "cylindrical" ;
        byte DT_analysis (time, nj, ni) ;
                DT_analysis:long_name = "deviation from sst reference climatology" ;
                DT_analysis:units = "kelvin" ;
                DT_analysis:_FillValue = -128 ;
                DT_analysis:add_offset = 0. ;
                DT_analysis:scale_factor = 0.1 ;
                DT_analysis:valid_min = -127 ;
                DT_analysis:valid_max = 127 ;
                DT_analysis:coordinates = "lon lat" ;
                DT_analysis:grid_mapping = "cylindrical" ;
                DT_analysis:reference = "climatology, Faugere and all" ;
```

```
byte surface_solar_irradiance(time, nj, ni) ;
        surface_solar_irradiance:long_name = "surface solar irradiance" ;
        surface_solar_irradiance:units = "W m-2" ;
        surface_solar_irradiance:_FillValue = -128 ;
        surface_solar_irradiance:add_offset = 250. ;
        surface_solar_irradiance:scale_factor = 0.2 ;
        surface_solar_irradiance:valid_min = -127 ;
        surface_solar_irradiance:valid_max = 127 ;
        surface_solar_irradiance:coordinates = "lon lat" ;
        surface_solar_irradiance:grid_mapping = "cylindrical" ;
byte ssi_dtime_from_sst(time, nj, ni) ;
        ssi_dtime_from_sst:long_name = "time difference of surface solar irradiance measurement from sst measurement" ;
        ssi_dtime_from_sst:units = "hour" ;
         ssi_dtime_from_sst:_FillValue = -128 ;
        ssi_dtime_from_sst:add_offset = 0. ;
        ssi_dtime_from_sst:scale_factor = 0.1 ;
        ssi_dtime_from_sst:valid_min = -127 ;
        ssi_dtime_from_sst:valid_max = 127 ;
        ssi_dtime_from_sst:coordinates = "lon lat" ;
        ssi_dtime_from_sst:grid_mapping = "cylindrical" ;
byte wind_speed(time, nj, ni) ;
        wind_speed:long_name = "wind speed" ;
        wind_speed:units = "m s-1" ;
        wind_speed:_FillValue = -128 ;
        wind_speed:add_offset = 0. ;
        wind_speed:scale_factor = 1. ;
        wind_speed:valid_min = -127 ;
        wind_speed:valid_max = 127 ;
        wind_speed:coordinates = "lon lat" ;
        wind_speed:grid_mapping = "cylindrical" ;
byte wspd_dtime_from_sst(time, nj, ni) ;
        wspd_dtime_from_sst:long_name = "time difference of wind speed measurement from sst measurement" ;
        wspd_dtime_from_sst:units = "hour" ;
        wspd_dtime_from_sst:_FillValue = -128 ;
        wspd_dtime_from_sst:add_offset = 0. ;
        wspd_dtime_from_sst:scale_factor = 0.1 ;
        wspd_dtime_from_sst:valid_min = -127 ;
        wspd_dtime_from_sst:valid_max = 127 ;
        wspd_dtime_from_sst:coordinates = "lon lat" ;
        wspd_dtime_from_sst:grid_mapping = "cylindrical" ;
byte sea_ice_fraction(time, nj, ni) ;
        sea_ice_fraction:long_name = "sea ice fraction" ;
        sea_ice_fraction:units = "percent" ;
        sea_ice_fraction:_FillValue = -128 ;
         sea_ice_fraction:add_offset = 0. ;
        sea_ice_fraction:scale_factor = 0.01 ;
        sea_ice_fraction:valid_min = 0 ;
        sea_ice_fraction:valid_max = 100 ;
        sea_ice_fraction:coordinates = "lon lat" ;
        sea_ice_fraction:grid_mapping = "cylindrical" ;
byte aerosol_optical_depth(time, nj, ni) ;
        aerosol_optical_depth:long_name = "aerosol optical depth" ;
        aerosol_optical_depth:units = "count'" ;
        aerosol_optical_depth:_FillValue = -128 ;
        aerosol_optical_depth:add_offset = 0. ;
         aerosol_optical_depth:scale_factor = 1. ;
        aerosol_optical_depth:valid_min = -127 ;
        aerosol_optical_depth:valid_max = 127 ;
         aerosol_optical_depth:coordinates = "lon lat" ;
        aerosol_optical_depth:grid_mapping = "cylindrical" ;
byte sources_of_ssi(time, nj, ni) ;
        sources_of_ssi:long_name = "sources_of surface solar irradiance" ;
        sources_of_ssi:_FillValue = -128 ;
        sources_of_ssi:comment = "details here source codes" ;
        sources_of_ssi:coordinates = "lon lat" ;
        sources_of_ssi:grid_mapping = "cylindrical" ;
byte sources_of_sea_ice_fraction (time, nj, ni) ;
        sources_of_sea_ice_fraction:long_name = "sources of sea ice fraction " ;
        sources_of_sea_ice_fraction:_FillValue = -128 ;
        sources_of_sea_ice_fraction:comment = "details here source codes" ;
        sources_of_sea_ice_fraction:coordinates = "lon lat" ;
        sources_of_sea_ice_fraction:grid_mapping = "cylindrical" ;
byte sources_of_aod(time, nj, ni) ;
         sources_of_aod:long_name = "sources of aerosol optical depth " ;
```

```
        sources_of_aod:_FillValue = -128 ;
        sources_of_aod:comment = "details here source codes" ;
        sources_of_aod:coordinates = "lon lat" ;
        sources_of_aod:grid_mapping = "cylindrical" ;
byte satellite_zenith_angle(time, nj, ni) ;
        satellite_zenith_angle:long_name = "satellite zenith angle" ;
        satellite_zenith_angle:units = "angular_degree" ;
        satellite_zenith_angle:_FillValue = -128 ;
        satellite_zenith_angle:add_offset = 0. ;
        satellite_zenith_angle:scale_factor = 1. ;
        satellite_zenith_angle:valid_min = -90 ;
        satellite_zenith_angle:valid_max = 90 ;
        satellite_zenith_angle:coordinates = "lon lat" ;
        satellite_zenith_angle:grid_mapping = "cylindrical" ;
byte rejection_flag(time, nj, ni) ;
        rejection_flag:long_name = "rejection flag" ;
        rejection_flag:comment = "
                b0 : 1= SST out of range;
                b1 : 1= Cosmetic value;
                b2 : 1= IR cloudy
                b3 : 1= MW rain
                b4 : 1= MW ice
                b5 : 1= MW wind
                b6 : 1= Land;
                b7 : 1= spare " ;
        rejection_flag:coordinates = "lon lat" ;
        rejection_flag:grid_mapping = "cylindrical" ;
byte confidence_flag(time, nj, ni) ;
        confidence_flag:long_name = "confidence flag" ;
        confidence_flag:comment = "
                b0 : 1=potential side lobe contamination;
                b1 : 1=relaxed rain contamination suspected;
                b2 : 1=TMI SST retrieved in SST < 285K
                b3 : 1=high wind speed retrieval
                b4 : 1=sea ice retrieval for MW data
                b5 : 1= sun glint suspected
                b6 : 1= L2 native bias and standard deviation;
                b7 : 1= L2 native confidence value " ;
        confidence_flag:coordinates = "lon lat" ;
        confidence_flag:grid_mapping = "cylindrical" ;
byte proximity_confidence_value(time, nj, ni) ;
        proximity_confidence_value:long_name = "proximity confidence value" ;
        proximity_confidence_value:_FillValue = -128 ;
        proximity_confidence_value:coordinates = "lon lat" ;
        proximity_confidence_value:grid_mapping = "cylindrical" ;

// global attributes:
Conventions = "CF-1.0" ;
:title = "Sea Surface Temperature from AVHRR onboard NOAA-16, resampled to 0.01 degrees resolution, over Mediterranean" ;
:DSD_entry_id = "EUR-CMS-AVHRR16_L-IR-HRDDS-MOCC" ;
:location = "SIM007"
:references = "Medspiration products user manual, Robinson I., Leborgne P., Piolle J.F., Larnicol G., v1.02, September 2004" ;
:institution = "MEDSPIRATION" ;
:contact = "isr@soc.soton.ac.uk" ;
:GDS_version_id = "v1.0-rev1.4" ;
:necdf_version_id = "3.5" ;
:creation_date = "2004-08-25" ;
:product_version = "1.0" ;
:history = " " ;
:platform = "NOAA-16" ;
:sensor = "avhrr" ;
:spatial_resolution = "2 km;
:start_date = "2004-08-25 UTC" ;
:start_time = "00:12:23 UTC" ;
:stop_date = "2004-08-25 UTC" ;
:stop_time = "00:14:18 UTC" ;
:southernmost_latitude = "43.37" ;
:northernmost_latitude = "45.37f" ;
:westernmost_longitude = "7.90f" ;
:easternmost_longitude = "9.90f" ;
:file_quality_index = "0" ;
:comment = " " ;
}
```

A5.3 HR-DDS metadata record format and registration

HR-DDS granule metadata records are identical in content and format to GDS MMR metadata records. A full description of MMR-DSD and MMR_FR is given in Appendix A6.

A5.4 Location of HR-DDS sites

Tables A5.4.1, A5.4.2, A5.4.3 and A5.4.4 describe the locations and formal names of each HR-DDS (v2) site for open ocean, moored buoys, SIMBIOS and other sites respectively.

Table A5.4.1 GODAE/GHRSST-PP Diagnostic Data Set (DDS) Open Ocean coordinates v2.3 (April 2003) Based on output of the 2nd & 3rd GHRSST-PP workshop Science Team feedback.

HRDDS Long Name	HRDDS Short Name	Latitude (lower left)	Longitude (lower left)	Δy-size	Δx-size	Description	Comment
GHRSST- Madeira	ghr001	35	-15	2.0	2.0	Atlantic Ocean Madeira	
GHRSST-North-NE-Atlantic	ghr002	50	-30	2.0	2.0	N Atlantic	
GHRSST-Central-NE-Atlantic	ghr003	30	-30	2.0	2.0	Central N Sub-tropical Atlantic	
GHRSST-Central-SE-Atlantic	ghr004	-20	-30	2.0	2.0	Central S Sub-tropical Atlantic	
GHRSST-Weddel-Sea-ICE	ghr005	-62.5	-30	25.0	5.0	Weddel Sea (ice edge)	
GHRSST-North-NW-Atlantic	ghr006	50	-45	2.0	2.0	W N Atlantic	
GHRSST-Central-NW-Atlantic	ghr007	30	-50	2.0	2.0	W Atlantic	
GHRSST-SW-Atlantic-Uruguay	ghr008	-38	-45	2.0	2.0	Uruguay coast	
GHRSST-North-Water-Pollyanna	ghr009	76.5	-74	2.0	2.0	North Water Pollyanna	
GHRSST-Drake-Passage-ICE	ghr010	-62.5	-70	15.0	5.0	Drake passage (ice edge)	
GHRSST-Chile-Coquimbo	ghr011	-30	-74	2.0	2.0	Chile	
GHRSST-South-SE-Pacific	ghr012	-45	-90	2.0	2.0	Pacific-Antarctic basin	
GHRSST-Easter_Island	ghr013	-25	-105	2.0	2.0	Easter Island	
GHRSST-South-Pacific-Antarctic-Ridge	ghr014	-45	-110	2.0	2.0	Pacific Antarctic Ridge	
GHRSST-Amudsen-Sea-ICE	ghr015	-67.5	-110	15.0	5.0	Amudsen Sea (ice edge)	
GHRSST-Pitcairn-Island	ghr016	-37	-130	2.0	2.0	Pitcairn Island S Pacific	
GHRSST-Central-N-Pacific	ghr017	30	-150	2.0	2.0	Central N Pacific	
GHRSST-Central-S-Pacific	ghr018	-37	-150	2.0	2.0	S Central Pacific	
GHRSST-W-Ross-Sea-ICE	ghr019	-65	-150	20.0	5.0	W Ross Sea (ice edge)	
GHRSST-Chukchi-Sea	ghr020	70	-170	5.0	5.0	Chukchi Sea	
GHRSST-Aleutian-Islands	ghr021	40	-170	2.0	2.0	Aleutian Islands	
GHRSST-Kermadec-Trench	ghr022	-45	-170	2.0	2.0	Kermadec trench	S Pacific
GHRSST-W-Aleutian-Islands	ghr023	50	170	2.0	2.0	W Aleutian Islands	
GHRSST-North-NW-Pacific	ghr024	30	170	2.0	2.0	NW Pacific	
GHRSST-Norfolk-Island	ghr025	-30	170	2.0	2.0	Norfolk Is. S Pacific	
GHRSST-E-Ross-Sea-ICE	ghr026	-65	170	20.0	5.0	E Ross Sea (ice edge)	
GHRSST-Sea-of-Okhotsk	ghr027	55	150	2.0	2.0	Sea of Okhotsk	
GHRSST-Coral-Sea	ghr028	-15	150	2.0	2.0	Coral Sea	
GHRSST-S-Tasman-Ridge	ghr029	-45	150	2.0	2.0	S Tasman Ridge	
GHRSST-Southern-Ocean-ICE	ghr030	-61	130	12.0	5.0	S Ocean (ice edge)	
GHRSST-S-China-Sea	ghr031	7	110	2.0	2.0	S China Sea	
GHRSST-Java-Trench	ghr032	-15	110	2.0	2.0	Java Trench	
GHRSST-East-Southern-Ocean	ghr033	-45	110	2.0	2.0	Eastern S Ocean	

HRDDS Long Name	HRDDS Short Name	Latitude (lower left)	Longitude (lower left)	Δy-size	Δx-size	Description	Comment
GHRSST-Bay-Of-Bengal	ghr034	15	90	2.0	2.0	Bay of Bengal	
GHRSST-Central-Southern-Ocean	ghr035	-30	90	2.0	2.0	Central S. Ocean	
GHRSST-South-Central-Southern-Ocean	ghr036	-45	90	2.0	2.0	Central S. Ocean	
GHRSST	ghr037	-61	90	12.0	5.0	S Ocean (ice edge)	
GHRSST-Arabian-Sea	ghr038	20	65	2.0	2.0	Arabian Sea	
GHRSST-Somali-Jet	ghr039	10	55	2.0	2.0	Somali Jet	
GHRSST-Central-Indian-Ocean	ghr040	-10	70	2.0	2.0	Indian Ocean	
GHRSST-South-Indian-Ocean	ghr041	-30	70	2.0	2.0	S Indian Ocean	
GHRSST-Kerguelen-Island	ghr042	-45	70	2.0	2.0	Kerguelen Is. S Ocean	
GHRSST-Somali-Basin	ghr043	-5	50	2.0	2.0	Somali Basin	
GHRSST-Mauritius-Basin	ghr044	-30	50	2.0	2.0	Mauritius Basin	
GHRSST-SW-Indian-Ocean	ghr045	-45	50	2.0	2.0	S Indian Ocean	
GHRSST-Enderby-Land-ICE	ghr046	-61	50	12.0	5.0	S Indian Ocean Enderby Land (ice edge)	
GHRSST-Agulhas-Basin	ghr047	-45	30	2.0	2.0	Agulhas Basin	
GHRSST-Ionian-Sea	ghr048	36	19	2.0	2.0	Mediterranean Ionian Sea	
GHRSST-West-Tropical-Atlantic	ghr049	-30	0	2.0	2.0	SE Atlantic	
GHRSST-Dronning-Maud-Land-ICE	ghr050	-60	10	20.0	5.0	SE Atlantic (ice edge)	
GHRSST-Davis-Strait	ghr051	60.0	-55.0	5.0	5.0	Davis Strait	
GHRSST-Denmark-Strait	ghr052	60.0	-30.0	2.0	2.0	Denmark Strait	
GHRSST-Black-Sea	ghr053	43.0	31.5	2.0	2.0	Black Sea	
GHRSST-Caspian-Sea	ghr054	42.5	50.0	2.0	2.0	Caspian Sea	
GHRSST-Lake-Tahoe	ghr055	39.2	-120.0	1.0	1.0	Lake Tahoe	USA
GHRSST-Quinghai-Hu-Lake	ghr056	36.8	100.2	1.0	1.5	Qinghai Hu Lake	China
GHRSST-Red-Sea	ghr057	17.5	40.0	2.0	2.0	Red Sea	
GHRSST-Peru-Current	ghr058	-15.0	-80.0	2.0	2.0	Peru current	
GHRSST-IMET-Peru-Stratus	ghr059	-20.0	-85.0	2.0	2.0	IMET/EPIC mooring under stratus cloud	
GHRSST-Western-Med	ghr060	40.0	6.0	2.0	2.0	Western Mediterranean	
GHRSST-FantaseaFerry-AIMS	ghr061	-20.0	149.0	2.0	2.0	Great Barrier Reef, Fantasea Ferry, AIMS	
GHRSST-Hudson-Bay	ghr062	59.0	-86.0	2.0	2.0	Hudson Bay	
GHRSST-Greenland-Sea-ICE	ghr063	75.0	0.0	12.0	5.0	Greenland Sea and ice edge	
GHRSST-Florida-Keys	ghr064	25.0	-81.0	2.0	2.0	Florida Keys (C-MAN sites)	
GHRSST-Central-S-Atlantic	ghr065	-40.0	-20.0	2.0	2.0	S Atlantic	
GHRSST-Barents-Sea	ghr066	75.0	40.0	5.0	5.0	Barents Sea	
GHRSST-Laptev-Sea	ghr067	74.0	130.0	5.0	5.0	Laptev Sea	
GHRSST-Tuamotu-Pacific	ghr068	-20.0	-140.0	2.0	2.0	South Pacific	
GHRSST-Perth	ghr069	-32.0	115.0	2.0	2.0	Perth-Rottnest Island Ferry	

Table A5.4.2 GODAE/GHRSST-PP Diagnostic Data Set (DDS) Moored buoy site coordinates v2.3 (April 2003) Based on output of the 2nd & 3rd GHRSST-PP workshop Science Team feedback.

Name	GHRSST short name)	Latitude (lower left)	Longitude (lower left)	Δy-size	Δx-size	Description	Comment
NDBC-46035	ndb051	57.08	-177.71	2.0	2.0	Bering Sea	
TAO38	tao038	8.0	-170.0	2.0	2.0		(Aug 1992)
TAO44	tao044	-8.0	-170.0	2.0	2.0		(Aug 1992)
TAO41	tao041	0.0	-170.0	2.0	2.0		(May 1998)
NDBC-46066	ndb063	52.65	-155.00	2.0	2.0	Kodiak	
TAO31	tao031	8.0	-155.0	2.0	2.0		(Aug 1992)
TAO37	tao037	-8.0	-155.0	2.0	2.0		(Mar 1992)
NDBC-46006	ndb039	40.84	-137.49	2.0	2.0	SW. Astoria	
TAO18	tao018	8.0	-125.0	2.0	2.0		(Oct 1992)
TAO21	tao021	0.0	-125.0	2.0	2.0		(Oct 1983)
TAO24	tao024	-8.0	-125.0	2.0	2.0		(Sep 1992)
TAO25	tao025	9.0	-140.0	2.0	2.0		(May 1988)
TAO28	tao028	0.0	-140.0	2.0	2.0		(Apr 1983)
TAO30	tao030	-5.0	-140.0	2.0	2.0		(Oct 1990)
TAO11	tao011	8.0	-110.0	2.0	2.0		(Oct 1991)
TAO14	tao014	0.0	-110.0	2.0	2.0		(Jan 1979)
TAO17	tao017	-8.0	-110.0	2.0	2.0		(Nov 1985)
TAO1	tao001	12.0	-95.0	2.0	2.0		(Dec 1999)
TAO7	tao007	0.0	-95.0	2.0	2.0		(Jul 1981 inactive 83-92)
TAO10	tao010	-8.0	-95.0	2.0	2.0		(Aug 1994)
CAN-46147	can001	51.83	-131.22	2.0	2.0	Vancouver	
MF-41100	mfr001	15.9	-57.9	2.0	2.0	Antilles (Meteo France)	
CAN-44141	can002	42.10	-56.22	2.0	2.0	Gulf Stream	
EGOS-62081	ego003	51.0	-13.3	2.0	2.0	Celtic Sea	
PIRATA-Reggae	pir007	15.0	-38.0	2.0	2.0		(Jan 1998)
PIRATA-Forro	pir008	12.0	-38.0	2.0	2.0		(Feb 1999)
PIRATA-Lambada	pir009	8.0	-38.0	2.0	2.0		(Jan 1998)
PIRATA-Frevo	pir010	4.0	-38.0	2.0	2.0		(Feb 1999)
PIRATA-Samba	pir006	0.0	-35.0	2.0	2.0		(Jan 1998)
PIRATA-Jazz	pir005	0.0	-23.0	2.0	2.0		(Mar 1999)
PIRATA-Soul	pir001	0.0	0.0	2.0	2.0		(Feb 1998)
PIRATA-Java	pir002	0.0	-10.0	2.0	2.0		(Sep 1997)
PIRATA-Valse	pir003	-6.0	-10.0	2.0	2.0		(Mar 2000)
PIRATA-Gavotte	pir004	-10.0	-10.0	2.0	2.0		(Sep 1997)
UKMO-63117	ukm004	58.0	1.10	2.0	2.0	North Sea	
TAO/TRITON	tao073	-5.0	95.0	2.0	2.0		(26 Oct 2001)
NDBC-45001	ndb027	48.06	-87.78	2.0	2.0	Mid Lake Superior	USA

Name	GHRSST short name)	Latitude (lower left)	Longitude (lower left)	Δy-size	Δx-size	Description	Comment
TAO/TRITON	tao069	8.0	137.0	2.0	2.0		(28 Sep 2001)
TAO/TRITON	tao071	0.0	138.0	2.0	2.0		(03 Oct 2001)
TAO/TRITON	tao065	5.0	147.0	2.0	2.0		(Feb 1990; Triton from 1999)
TAO/TRITON	tao067	0.0	147.0	2.0	2.0		(Apr 1994; Triton from 1999)
TAO/TRITON	tao059	8.0	156.0	2.0	2.0		(Dec 1994; Triton from 1999)
TAO/TRITON	tao062	0.0	156.0	2.0	2.0		(Jul 1995; Triton from 1999)
TAO/TRITON	tao064	-5.0	156.0	2.0	2.0		(Aug 1991; Triton from 1999)
TAO55	tao055	0.0	165.0	2.0	2.0		(Jan 1986)
TAO52	tao052	8.0	165.0	2.0	2.0		(Jul 1989)
TAO58	tao058	-8.0	165.0	2.0	2.0		(Aug 1991)
TAO45	tao045	8.0	180.0	2.0	2.0		(Nov 1993)
TAO48	tao048	0.0	180.0	2.0	2.0		(Mar 1993)
TAO51	tao051	-8.0	180.0	2.0	2.0		(Nov 1993)
NDBC-42002	ndb008	25.17	-94.42	2.0	2.0	Western Gulf	
GHRSST-Nauru	ghrsst69	-0.32	161.55	2.0	2.0	Nauru Island	
GHRSST-J-MOOR	ghrsst70	28	134	2.0	2.0	Japanese long term mooring site with ~50 years of data	
GHRSST-Barrow	ghrsst71	71.22	160.3	2.0	2.0	Barrow point, Alaska	
GHRSST-Manaus	ghrsst72	-3.0	146.0	2.0	2.0	Manaus island	
GHRSST-ISAR-validation	ghrsst73	43	7	8	-8	Bay of Biscay and E. Channel area for ISAR validation transects	May 2004-2006
GHRSST-MAERI-validation	ghrsst74	12	88	5	-20	Track of the Royal Carribbean Explorere of the seas around Cuba	
GHRSST-POL-Irish-sea	ghrsst75	52	6.5	3	-4	Location of the Proudman Oceanographic Laboratory POLCOMS model and POL Coatstal observatory system	
GHRSST-shelf-seas	ghrsst76	50	2	7	-9	North sea area for comparison with MEtOFfice shelf seas operational model output	

Table A5.4.3 GODAE/GHRSST-PP Diagnostic Data Set (DDS) SIMBIOS Diagnostic Data Set Site coordinates http://simbios.gsfc.nasa.gov/Info/STM2001/Sites.html .

Name	GHRSST short name)	Latitude (lower left)	Longitude (lower left)	Δy-size	Δx-size	Description	Comment
SIMBIOS-MOBY	sim001	20.8	-157.20	2.0	2.0	Hawaii	
SIMBIOS-BATS/BTM	sim002	32.0	-64.50	2.0	2.0	Bermuda	
SIMBIOS-CALCOFI	sim003	29.85	-123.59	2.0	2.0	California	
SIMBIOS-EqPAC	sim004	0.0	-155.00	2.0	2.0	Eastern Equ Pacific	
SIMBIOS-HOT	sim005	22.75	-158.00	2.0	2.0	Hawaii	
SIMBIOS-Ishigaki	sim006	24.39	123.27	2.0	2.0	East China Sea	
SIMBIOS-Ligurian_Sea	sim007	43.37	7.90	2.0	2.0	Mediterranean	
SIMBIOS-Lower_Chesapeake_Bay	sim008	37.40	-76.13	2.0	2.0	Virginia	
SIMBIOS-Monterey_Bay	sim009	36.75	-122.42	2.0	2.0	Monterey	
SIMBIOS-Plymbody	sim010	50.2	-4.10	2.0	2.0	English Channel	
SIMBIOS-Venice_Tower	sim011	45.31	12.60	2.0	2.0	Northern Adriatic	
SIMBIOS-Station_H	sim012	41.5	145.78	2.0	2.0	Japan East Coast	
SIMBIOS-Cariaco_Basin	sim013	10.5	-64.66	2.0	2.0	Venezuela	
SIMBIOS-Kashidoo	sim014	4.95	73.45	2.0	2.0	Maldives Islands	
SIMBIOS-Korean	sim015	32.0	125.00	2.0	2.0	East China Sea	
SIMBIOS-LEO_15	sim016	39.3	-74.25	2.0	2.0	New Jersey	
SIMBIOS-Plumes_and_Blumes	sim017	34.25	-119.92	2.0	2.0	off Santa Barbara CA	
SIMBIOS-Scotian_Prince_Route	sim018	43.00	-69.00	2.0	2.0	Gulf of Maine	
SIMBIOS-NOAA-GOM	sim019	29.50	-87.50	2.0	2.0	Northern Gulf of Mexico	
SIMBIOS-NOAA-NC	sim020	35.00	-76.50	2.0	2.0	off North Carolina	
SIMBIOS-Rottnest_Island	sim021	-31.80	115.30	2.0	2.0	off Western Australia	

Table A5.4.4 GODAE/GHRSST-PP Diagnostic Data Set (DDS) general site coordinates v2.3 (April 2003) Based on output of the 2nd & 3rd GHRSST-PP workshop Science Team feedback.

Location (GCMD location_valid)	GHRSST short name)	Latitude (lower left)	Longitude (lower left)	Δy-size	Δx-size	Description	Comment
SIMBIOS-Arm_1	sim022	0.00	168.00	2.0	2.0	Nauru Island	
SIMBIOS-Arm_2	sim023	25.00	148.00	2.0	2.0	Manus Island	
SIMBIOS-NW_Afr_Upwell	sim025	21.00	-17.50	2.0	2.0	Morocco	
SIMBIOS-Alberon_Gyre	sim026	33.00	32.50	2.0	2.0	Eastern Mediterranean	
SIMBIOS-Helgoland	sim027	54.00	9.00	2.0	2.0	North Sea	
SIMBIOS-Nordic	sim028	55.00	19.30	2.0	2.0	Baltic Sea	
SIMBIOS-Luderitz_Upwell	sim029	-26.00	14.50	2.0	2.0	Namibia	
SIMBIOS-Philippine_Sea	sim030	17.00	133.00	2.0	2.0	Southeast Asia	
SIMBIOS-Cook_Islands	sim031	-20.00	-163.00	2.0	2.0	New Zealand	

Appendix A6. GHRSST-PP GDS Master Metadata Repository (MMR)

In order to successfully exchange and manipulate data in a real time environment, RDAC and GDAC must know at any instant what data are available, where they are and what level of processing they have completed. This is a particularly demanding task given the global distribution of RDAC and GDAC that are working together in the GHRSST-PP Regional/Global Task sharing framework. In order to address these issues, a global GHRSST-PP Master Metadata Repository (MMR) service has been developed.

The MMR ensures that all GDS data products are "visible" to the RDAC and GDAC centres at any given time. From a user perspective, the MMR provides a simple searchable index to all GHRSST-PP data products, irrespective of where they are physically located. The MMR is therefore the foundation of the GHRSST-PP data management framework. It provides a searchable catalogue of the distributed GHRSST-PP data holdings providing information on its physical location, contents and any constraints on its use. Using the results of a search submitted to the MMR system, a user may access data sets directly via ftp/OPeNDAP/LAS. Without a MMR master catalogue, it would be extremely difficult (if not impossible) to locate a single data resource within the GHRSST-PP without physically connecting to the computers storing the data and searching each one individually.

This appendix describes the GHRSST-PP Master Metadata Repository (MMR) service, the format of GHRSST-PP metadata records that are generated for each GDS data file and the mechanisms adopted for the delivery registration, modification and retrieval from the MMR system.

A6.1 The GHRSST-PP Master Metadata Repository (MMR) system

The GHRSST-PP data files themselves have been chosen to follow the Climate and Forecast netCDF conventions (See Appendix A1) because these conventions provide a practical standard for storing oceanographic data, and have already been adopted for the Data Sharing Pilot project within GODAE. The global attributes of netCDF files are also metadata, since they follow the CF conventions, although they differ slightly from the GHRSST-PP MMR standard. To a certain extent, this is reasonable, since the MMR is intended to help users and processing centres search for data sets and to keep track of processing tasks, while the netCDF attributes are to help users apply the data.

The following section describes the format of GHRSST-PP metadata records.

A6.2 Implementation of the GHRSST-PP MMR

The MMR will be implemented as a physical relational database (MySQL) that will contain metadata descriptions of all static (i.e., data products) data holdings within the GHRSST-PP. Metadata describing each data product including will be stored within MMR database system in real time. The GHRSST-PP MMR data record format is based on the Global Change Master Directory (GCMD) Directory Interchange Format (DIF) standard. The GCMD has an extensive number of keyword lists that can be used for encoding the physical quantities, the data centres, *etc*. Use of these keyword lists provides a simple and efficient way for users to make unambiguous searches of the database in order to locate the data they want. GCMD also has a large set of sensibly defined metadata fields based on many years of experience of serving metadata. Metadata records will be created and submitted to the MMR system as XML coded records according to specific DTD specifications within the GDS.

The design of the GHRSST-PP "master" metadata repository (MMR) is based on a relational database (MySQL) that is indexed and searchable through a web interface front end (see http://www.ghrsst-pp.org) communicating with a back end database server (MySQL). The database contents will be accessed through Perl and SQL function calls with results formatted in XML that can be displayed in a web browser. The query results can also be formatted into DODS or FTP syntax to allow immediate and direct access to the DDS where the data resides (a location different than the metadata repository). The goal is to keep the metadata storage and query on a very simple level while maintaining good functionality with possibility for future modification and expansion.

An automated metadata ingest system will be available (see Figure A6.2.1) to ingest metadata submissions as new data records are added (including high resolution, regional and global). This ingest system will be in the simple form of email submissions emanating from the RDAC/GDAC that are parsed with scripts to extract relevant metadata content before the database is populated. A second interface will allow RDAC to push MMR records via ftp that will be then automatically ingested by the MMR system.

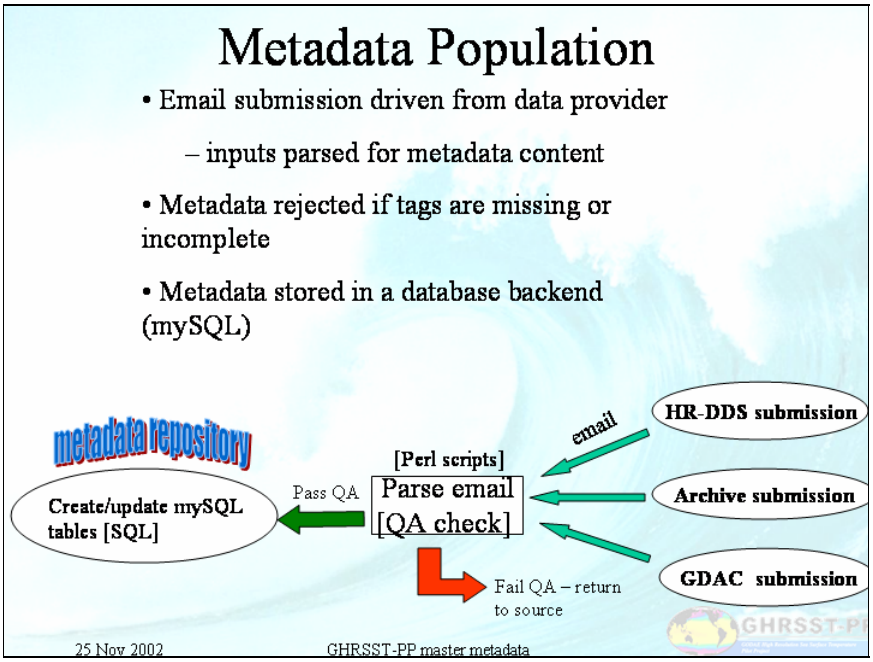


Figure A6.2.1. Proposed automatic system for the GHRSST-PP Master Metadata Repository system. (E. Armstrong).

A quality control system will be implemented to determine if the metadata adheres to the proper format (See Appendix A6.2.2). The submitting individual will be responsible for formatting the metadata input correctly according to the required fields and metadata templates. Metadata creation tools and methodology will be explored to assist the submission process.

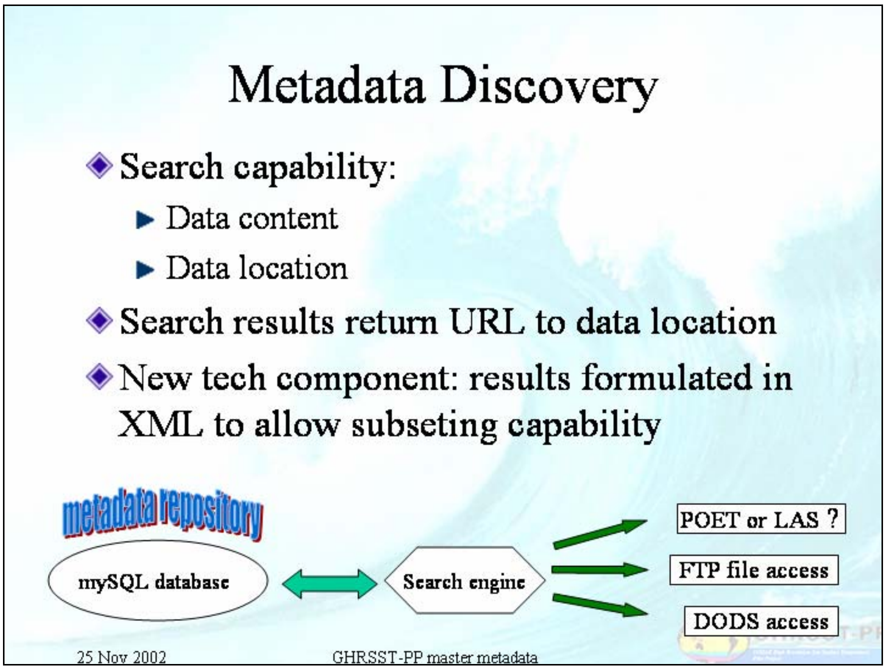


Figure A6.2.1. MMR system for GHRSST-PP, data discovery (E. Armstrong)

In order to make use of the MMR, a data discovery system will be developed that allows a user to search for data by several fields addressing the data content (at a minimum by time and location). The search should then return a URL to the location of the data within the GHRSST-PP distributed data set (see Figure 4.2). It is proposed that XML is used for reporting results so that sub-setting of results can be undertaken.

There are a variety of options available for sub-setting data including the PO.DAAC in house POET system, the Live Access Server (LAS) or via DODS/OPeNDAP access. The MMR will be designed with each of these systems in mind.

A6.3 GHRSST-PP Master Metadata Repository metadata record format

The metadata standard defined for GHRSST-PP MMR system is based on the GCMD DIF standard. The following web reference documents are applicable to GDS metadata records:

GCMD DIF definition document http://gcmd.gsfc.nasa.gov

GHRSST-PP metadata consist of two related but distinct types of metadata record:

1. A Data Set Description (MMR-DSD) contains information describing a **data set** defined as a collection of data files that have generic contents. One and only one DSD record exists for every data set used within the GHRSST-PP and contains information that is common to all of the data files that collectively consititue a data set (e.g., Satellite sensor, contact information). An MMR_DSD may be thought of as a cover page and an abstract for a data set. In addition, a DSD exists for every GHRSST-PP High Resolution Diagnostic Data Set (HR-DDS) site as explained in Appendix A5. The format of MMR-DSD records are given in Table A6.3.1.
2. Each file in a data set described by a MMR-DSD is represented by a unique File Record (MMR_FR) metadata record. A MMR_FR metadata record contains information that distinguishes a specific data file from all others in a given data set. The format of MMR_FR is given in Table A6.3.2. There are many MMR_FR metadata records associated with a single MMR-DSD metadata record.

Within the GDS only MMR_FR metadata records need to be created and submitted to the MMR system operationally by RDAC and GDAC. DSD metadata records will be created by the GHRSST-PO in collaboration with data providers, RDAC and GDAC. Every time a GDS data file or is produced, an associated MMR_FR metadata record must be created (see Appendix A6.4) and registered at the MMR via e-mail (see Section A6.5). Figure 6.3.1 summarises the relationship between MMR_DSD and MMR_FR.

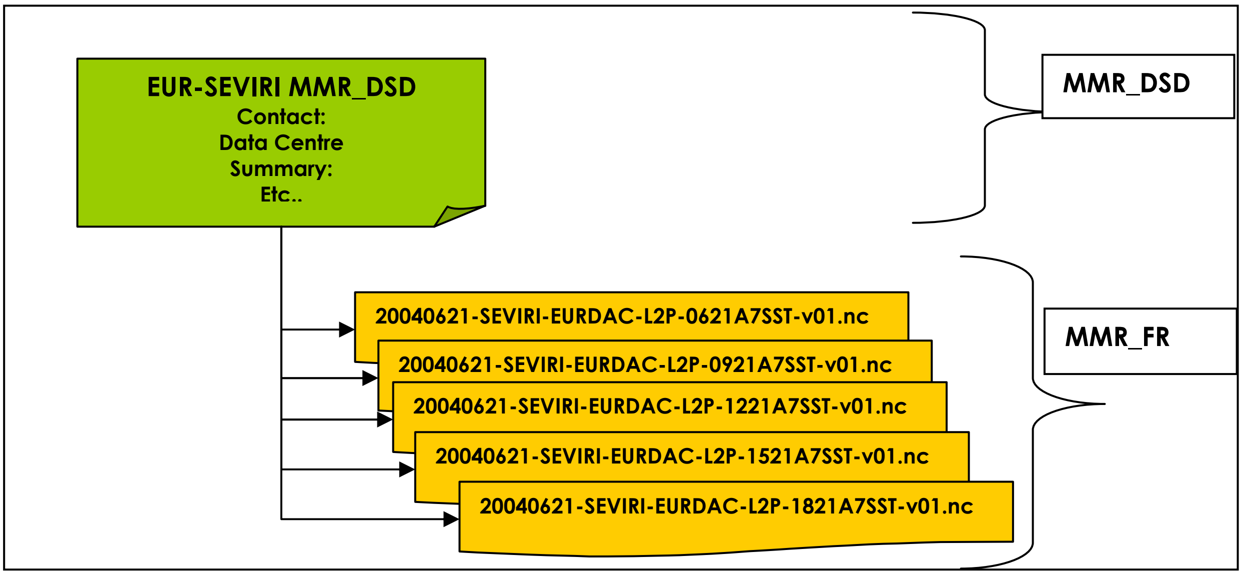


Figure 6.3.1 Schematic diagram showing the relationship between MMR_DSD (one per data set) and MMR_FR (many per data set)

All GHRSST-PP MMR metadata records are encoded using the extensible mark-up language (XML). XML provides an application independent way of sharing and controlling the quality and content of data. A detailed description of XML-v1.0 can be found at http://www.w3.org/TR/REC-xml

An XML Document Type Definition (DTD) is a formal description in XML Declaration Syntax of a particular type of document. A DTD provides applications with advance notice of what names and structures can be used in a particular document type. Using a DTD when editing files means you can be certain that all documents which belong to a particular type will be constructed and named in a consistent and conformant manner. Validating XML parsers read a DTD before they read an XML document so that they can identify where every element type ought to come and how each element relates to the other. A major advantage if using an XML DTD is that data records can be validated providing a quality control method well suited to database applications such as the MMR that are expecting a data record in a specific format.

A DTD can be declared inline within an XML document itself, or as an external reference. The MMR uses a GHRSST-PP XML DTD to verify that MMR metadata records received from the GHRSST-PP RDAC and GDAC are valid. MMR DTD can also be used to validate MMR data records at the time of metadata record creation.

A6.3.1 MMR_DSD metadata record format

A6.3.1.1 MMR_DSD file name convention

The MMR-DSD files for L2P products shall be named as follows:

```
DSD-<L0 ID>-<Processing Centre Code>-L2P-<L2 product type>-<Processing Model ID>.xml
```

Refer to Table A1.2.1 (L2P file naming conventions) for the meaning of each code. For example:

DSD-AVHRR16_L-EUR-L2P-LAC-v01.xml

The MMR-DSD files for L4/UHR products shall be named as follows:

DSD-<Processing Centre Code>-L4<Product type>-<Area>-<Processing Model ID>.xml

Refer to Table A1.3.1 (UHR/L4 file naming conventions) for the meaning of each code. For example:

DSD-EUR-L4UHfnd-MED-v01.xml

A6.3.1.2 MMR_DSD XML document type definition (DTD)

MMR-DSD files shall be formatted in XML format, complying to the following DTD:

```
<!-- DTD for the GHRSST-PP MMR_DSD                          -->
<!--                                                        -->
<!-- Author: J. F Piolle and C J Donlon                     -->
<!--                                                        -->
<!-- History                                                -->
<!-- 2004-02-17     Original                                -->
<!-- 2005-01-11     Revised time  formats and character lengths  -->


<!ELEMENT MMR_DSD ( Entry_ID,
                Entry_Title,
                Data_Set_Citation+,
                Parameters+,
                Sensor_Name+,
                Source_Name+,
                Temporal_Coverage,
                Spatial_Coverage,
                Location,
                Projection_Information,
                Data_Resolution,
                Access_Constraints,
                Use_Constraints,
                Originating_Center,
                Data_Center,
                Reference?,
                Summary,
                Related_URL*,
                DSD_File_Version,
                DSD_Creation_Date,
                Last_DSD_Revision_Date,
                DSD_Revision_History
               )
>

<!-- Entry ID and Title ========================================== -->
<!ELEMENT Entry_ID (#PCDATA)>
<!ELEMENT Entry_Title (#PCDATA)>

<!-- Data Set Citation ========================================== -->
<!ELEMENT Data_Set_Citation ( Dataset_Creator,
                        Dataset_Title,
                        Dataset_Series_Name?,
                        Dataset_Release_Date?,
                        Dataset_Release_Place?,
                        Dataset_Publisher?,
                        Version,
                        Other_Citation_Details?,
                        Online_Resource?
                       )
>

<!ELEMENT Dataset_Creator (#PCDATA)>
```

```
<!ELEMENT Dataset_Title (#PCDATA)>
<!ELEMENT Dataset_Series_Name (#PCDATA)>
<!ELEMENT Dataset_Release_Date (#PCDATA)>
<!ELEMENT Dataset_Release_Place (#PCDATA)>
<!ELEMENT Dataset_Publisher (#PCDATA)>
<!ELEMENT Version (#PCDATA)>
<!ELEMENT Issue_Identification (#PCDATA)>
<!ELEMENT Data_Presentation_Form (#PCDATA)>
<!ELEMENT Other_Citation_Details (#PCDATA)>
<!ELEMENT Online_Resource (#PCDATA)>

<!-- Parameters==================================================== -->
<!ELEMENT Parameters ( Category,
                   Topic,
                   Term,
                   Variable?,
                   Detailed_Variable?
                  )
>

<!ELEMENT Category (#PCDATA)>
<!ELEMENT Topic (#PCDATA)>
<!ELEMENT Term (#PCDATA)>
<!ELEMENT Variable (#PCDATA)>
<!ELEMENT Detailed_Variable (#PCDATA)>

<!-- Sensor Name ================================================== -->
<!ELEMENT Sensor_Name ( Short_Name,
                    Long_Name?
                   )
>

<!ELEMENT Short_Name (#PCDATA)>
<!ELEMENT Long_Name (#PCDATA)>

<!-- Source Name ================================================== -->
<!ELEMENT Source_Name ( Short_Name,
                    Long_Name?
                   )
>

<!-- Temporal Coverage ============================================ -->
<!ELEMENT Temporal_Coverage ( Start_Date?,
                          Stop_Date?,
                         )
>

<!ELEMENT Start_Date (#PCDATA)>
<!ELEMENT Stop_Date (#PCDATA)>


<!-- Spatial Coverage ============================================= -->
<!ELEMENT Spatial_Coverage ( Southernmost_Latitude,
                         Northernmost_Latitude,
                         Westernmost_Longitude,
                         Easternmost_Longitude,
                         Minimum_Altitude?,
                         Maximum_Altitude?,
                         Minimum_Depth?,
                         Maximum_Depth?
                        )
>

<!ELEMENT Southernmost_Latitude (#PCDATA) >
<!ELEMENT Northernmost_Latitude (#PCDATA) >
<!ELEMENT Westernmost_Longitude (#PCDATA) >
<!ELEMENT Easternmost_Longitude (#PCDATA) >
<!ELEMENT Minimum_Altitude (#PCDATA) >
<!ELEMENT Maximum_Altitude (#PCDATA) >
<!ELEMENT Minimum_Depth (#PCDATA) >
<!ELEMENT Maximum_Depth (#PCDATA)>

<!-- Location ===================================================== -->
<!ELEMENT Location ( Location_Name,
                 Detailed_Location?
                )
```

```
>

<!ELEMENT Location_Name (#PCDATA)>
<!ELEMENT Detailed_Location (#PCDATA)>

<!-- Projection information=========================================== -->
<!ELEMENT Projection_Information ( Projection_Type?,
                              Ellipsoid_Type?,
                              Other_Projection_Details?
                             )
>

<!ELEMENT Projection_Type (#PCDATA)>
<!ELEMENT Ellipsoid_Type (#PCDATA)>
<!ELEMENT Other_Projection_Details (#PCDATA)>

<!-- Data Resolution ============================================== -->
<!ELEMENT Data_Resolution ( Latitude_Resolution,
                         Longitude_Resolution,
                         Altitude_Resolution?,
                         Depth_Resolution?,
                         Temporal_Resolution
                        )
>

<!ELEMENT Latitude_Resolution (#PCDATA) >
<!ELEMENT Longitude_Resolution (#PCDATA) >
<!ELEMENT Altitude_Resolution (#PCDATA) >
<!ELEMENT Depth_Resolution (#PCDATA) >
<!ELEMENT Temporal_Resolution (#PCDATA) >

<!-- Access Constraints =========================================== -->
<!ELEMENT Access_Constraints (#PCDATA)>

<!-- Use Constraints ============================================== -->
<!ELEMENT Use_Constraints (#PCDATA)>

<!-- Originating Center =========================================== -->
<!ELEMENT Originating_Center (#PCDATA)>

<!-- Data Center ================================================== -->
<!ELEMENT Data_Center ( Data_Center_Name,
                      Data_Center_URL*,
                      Data_Set_ID?,
                      Personnel+
                     )
>

<!-- Data Center Name ============================================= -->
<!ELEMENT Data_Center_Name ( Short_Name,
                           Long_Name
                          )
>

<!ELEMENT Data_Center_URL (#PCDATA)>
<!ELEMENT Data_Set_ID (#PCDATA)>

<!-- Personnel ==================================================== -->
<!ELEMENT Personnel ( Role+,
                    First_Name,
                    Middle_Name?,
                    Last_Name,
                    Email,
                    Phone,
                    Fax?,
                    Address
                   )
>

<!ELEMENT Role (#PCDATA)>
<!ELEMENT First_Name (#PCDATA)>
<!ELEMENT Middle_Name (#PCDATA)>
<!ELEMENT Last_Name (#PCDATA)>
<!ELEMENT Email (#PCDATA)>
<!ELEMENT Phone (#PCDATA)>
<!ELEMENT Fax (#PCDATA)>
```

```
<!ELEMENT Address (#PCDATA)>

<!-- Reference ===================================================== -->
<!ELEMENT Reference (#PCDATA)>

<!-- Summary ====================================================== -->
<!ELEMENT Summary (#PCDATA)>

<!-- Related URL ================================================== -->
<!ELEMENT Related_URL ( URL_Content_Type?,
                        URL+,
                        Description?
                      )
>

<!ELEMENT URL_Content_Type (#PCDATA)>
<!ELEMENT URL (#PCDATA)>
<!ELEMENT Description (#PCDATA)>

<!-- DSD Creation Date ============================================ -->
<!ELEMENT DSD_File_Version (#PCDATA)>

<!-- DSD Creation Date ============================================ -->
<!ELEMENT DSD_Creation_Date (#PCDATA)>

<!-- Last DSD Revision Date ======================================= -->
<!ELEMENT DSD_Last_ Revision_Date (#PCDATA)>

<!-- DSD Revision History ========================================= -->
<!ELEMENT DSD_Revision_History (#PCDATA)>
```

A6.3.1.3 MMR_DSD XML DTD description

Table A6.1.3.3.1 provides a more explicit and readable description of the DTD. It is fully consistent with the DTD described in section A6.1.3.2 and should be use as reference.

Table A6.1.3.3.1 MMR_DSD XML description.

Field Name and attribute description	Cardinality	Type	Description
Entry_ID	1	Char(160)	A unique identifier for the data set. The identifier will be set by the project and reflect the location, source, sensor, parameter and product of the data set. Format should conform to the following scheme: <DataProcessingCentre[Table A2.1]>-<DataSetType> Where <DataSetType> is one of: L2 — Input data set as provided by data provider L2P — GHRSST-PP L2P data file L4SSTfnd — GHRSST-PP SSTfnd analysis data file L4UHRSSTfnd — Ultra-high resolution GHRSST-PP SSTfnd analysis data file NWP — Numerical Prediction system data file e.g., EURDAC-L4SSTfnd Refers to an analysed level-4 SST foundation data product file In addition, the following fields apply to L2P data: -<GHRSST-PP Identification Code [Tables A3.2.1, A3.3.1 and A3.5.1]> e.g., REMSS-L2P-AMSRE_SST Refers to an AMSR-E microwave L2P SST data file produced by Remote Sensing Systems)
Entry_Title	1	Char(256)	Data set title Title should be descriptive enough so that when a user is presented with a long list of titles, the user can determine the general content of the data set. In order to make titles descriptive, important elements about the data may be included in the title, i.e., parameters measured, geographic location, instrument, investigator, project, temporal coverage. For readability, capitalization of the title should follow standard constructs. Do not use all capital letters or all lower case letters, but use the appropriate case where applicable. Example "Sea Surface Temperature from the Along-Track Scanning Radiometer 2 (ATSR-2), 1km resolution, daily, global, for the year 1999, from the Rutherford Appleton Laboratory, UK"

Field Name and attribute description	Cardinality	Type	Description
Data_Set_Citation	1+		A citation for the data set to properly credit the data set producer. This field has two functions: to indicate how this data set should be cited in the professional scientific literature, and if this data set is a compilation of other data sets, to document and credit the data sets that were used in producing this compilation. This group is not to be used to list bibliographic references of scientific research articles arising from the data set. This field provides a citation for the data set itself, not articles related to the research results.
Dataset_Creator	1	Char(256)	The name of the organization(s) or individual(s) with primary intellectual responsibility for the data set's development
Dataset_Title	1	Char(256)	The title of the data set; this may be the same as Entry Title
Dataset_Series_Name	0 | 1	Char(160)	The name of the dataset series, or aggregate dataset of which the dataset is a part.
Dataset_Release_Date	0 | 1	Char(160)	The date when the data set was made available for release. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
Dataset_Release_Place	0 | 1	Char(100)	The name of the city (and state or province and country if needed) where the data set was made available for release
Dataset_Publisher	0 | 1	Char(100)	The name of the individual or organization that made the data set available for release
Version	1	Char(80)	The version of the data set
Other_Citation_Details	0 | 1	Char(256)	Any other details
Online_Resource	0 | 1	Char(256)	The URL of an online computer resource containing user-oriented information about the data set (e.g. help files, articles), if it exists.
Parameters	1+		This group describes the types of measurements represented by the data, such as sea surface temperature, wind speed, etc.
Category	1	Keyword[Char(31)]	These are the four GCMD DIF keyword fields: Category, Topic and Term, which cover a fairly exhaustive range of data types. Refer to : http://gcmd.gsfc.nasa.gov/Resources/valids/gcmd_parameters.html . Units should be stated in the data files themselves.
Topic	1	Keyword[Char(31)]	
Term	1	Keyword[Char(31)]	

Field Name and attribute description	Cardinality	Type	Description
Variable	1	Keyword[Char(31)]	
Detailed_Variable	0 | 1	Char(160)	A free text field for providing further specification of the measurement type
Sensor_Name	1+		The instrument or hardware used to acquire the data . Sensor Short_Name and Long_Name must selected in the DIF list if available (http://gcmd.gsfc.nasa.gov/Resources/valids/sensors.html). If the sensor does not appear in the Sensor Valids List, a new sensor name may be submitted to GCMD for review.
Short_Name	1	Keyword[Char(31)]	
Long_Name	0 | 1	Keyword[Char(81)]	
Source_Name	0+		The platform of data collection, as in a spacecraft, ship or ground station housing the sensor(s) used to acquire the data; or as in a map from which the data were digitized; or as in a model from which the data were generated. Source Short_Name and Long_Name must selected in the DIF list if available (http://gcmd.gsfc.nasa.gov/Resources/valids/sources.html). If the source does not appear in the Platform Valids List, a new name may be submitted to GCMD for review.
Short_Name	1	Keyword[Char(31)]	
Long_Name	0 | 1	Keyword[Char(81)]	
Temporal_Coverage	1		This group provides the start and end date and times for the measurements contained in the GHRSST-PP data file. All date and times must be in Universal Time Coordinates (UTC, which is approx. Greenwich Mean Time), and formatted at "yyyy-mm-dd" for dates and "hh:mm:ss" for times.
Start_Date	0 | 1	Char[17]	ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
Stop_Date	0 | 1	Char[17]	ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC

Field Name and attribute description	Cardinality	Type	Description
Spatial_Coverage	1		Geographic coverage (horizontal and vertical) of the data described. Longitude must be expressed in the range -180.0 to +360.0 as an offset from the Greenwich meridian, and latitude in the range -90.0 to +90.0, as an offset from the Equator. For gridded data, these coordinates should be the outer edges of the coverage, and not the coordinates of the centres of the edge pixels.
Southernmost_Latitude	1	Float	DD.ddd (2 decimals required)
Northernmost_Latitude	1	Float	DD.ddd (2 decimals required)
Westernmost_Longitude	1	Float	DDD.ddd (2 decimals required)
Easternmost_Longitude	1	Float	DDD.ddd (2 decimals required)
Minumum_Altitude	0 | 1	Char[80]	The altitude level specified in Meters which represents the lower limit of data coverage, as measured from mean sea level (Primarily for Model outputs)
Maximum_Altitude	0 | 1	Char[80]	
Minumum_Depth	0 | 1	Char[80]	The depth level specified in Meters which represents the upper-most depth of data coverage, as measured from mean sea level. (Primarily for Model outputs)
Maximum_Depth	0 | 1	Char[80]	
Location	1		Names of places which may be used for searching
Location_Name	1	keyword	Locations must be selected from the Location Valids (refer to http://gcmd.gsfc.nasa.gov/Resources/valids/location.html)
Detailed_Location	0 | 1	Char[80]	More specific names not specified in http://gcmd.gsfc.nasa.gov/Resources/valids/location.html may be added here if not in the location valids list and defining more precisely the location.
Projection_Information	1		
Projection_Type	0 | 1	Keyword[Char[160]]	A keyword from the GHRSST-PP approved list of projection names: (see table below)

Keyword	Description
Swath	Data in native swath format, may or may not be geo-referenced
Gridded	Data re-sampled onto a regular grid
Polar_Stereo	Polar stereographic
??	TBD as required

Field Name and attribute description	Cardinality	Type	Description
Ellipsoid_Type	0 | 1	Keyword[Char[160]]	A keyword from the GHRSST-PP approved list of ellipsoids[20]:

Keyword	Semi-major axis (meters)	1/Flattening
Airy_1830	6,377,563	299.33
Everest_1830	6,377,276.3	300.80
Bessel_1841	6,377,397.2	299.15
Clarke_1866	6,378,206.4	294.98
Clarke_1880	6,378,249.2	293.47
International_1924	6,378,388	297
Krasovsky_1940	6,378,245	298.3
IAU_1968	6,378,160	298.25
WGS_72	6,378,135	298.26
GRS_80	6,378,137	298.26
WGS_84	6,378,137	298.25722
User_Defined		

Field Name and attribute description	Cardinality	Type	Description
Other_Projection_Details	0 | 1	Char[1024]	Free text description of details such as the projections standard parallels, etc., if necessary to fully specify the projection.
Data_Resolution	1		
Latitude_Resolution	1	Char[80]	Spatial resolution specified in km
Longitude_Resolution	1	Char[80]	Spatial resolution specified in km
Altitude_Resolution	0 | 1	Char[80]	Specified in meters (Primarily for Model outputs)
Depth_Resolution	0 | 1	Char[80]	Specified in meters (Primarily for Model outputs)
Temporal_Resolution	0 | 1	Char[1024]	Text description of data resolution
Access_Constraints	1	Char[1024]	Restrictions, limitations and legal prerequisites for accessing the data set. Some words which may be used in this field include: Public, In-house, Limited, Source. e.g. "Data available to the general public after 5 year embargo period beginning January 1, 1994"
Use_Constraints	1	Char[1024]	Restrictions, limitations and legal prerequisites for using the data set. e.g. "Data may not be used for commercial applications".
Originating_Center	1	Char[160]	Data originating centre name (can be identical to Data_Center_Name but not necessarily)
Data_Center	1		This group identifies the data centre which distributes the data, the data centre URL, data set identification, and a person to contact. Note USA spelling of centre
Data_Center_Name	1		Name of the data centre taken from the list of valid keywords http://gcmd.gsfc.nasa.gov/Resources/valids/data_centre.html
Short_name	1	Keyword[Char[31]]	Name of the data centre taken from the list of valid keywords http://gcmd.gsfc.nasa.gov/Resources/valids/data_centre.html

Field Name and attribute description			Cardinality	Type	Description
		Long_name	1	Keyword [Char[160]]	Name of the data centre taken from the list of valid keywords http://gcmd.gsfc.nasa.gov/Resources/valids/data_centre.html
	Data_Center_URL		0+	Char[160]	The Internet Uniform Resource Locator(s) (URL) for the data centre should be listed if available. This field may be repeated as many times as necessary within the data centre group. The URLs will be hypertext linked. The URLs may include gopher, ftp and telnet as well as WWW servers e.g., Data_Center_URL: http://daac.gsfc.nasa.gov Data_Center_URL: ftp://podaac.jpl.nasa.gov Data_Center_URL: telnet://ncar.ucar.edu Data_Center_URL: gopher://www.ciesin.org
	Data_Set_ID		0|1	Char[160]	Should be listed if available. These are identification codes assigned by the data centre, which will simplify the location and ordering of the data sets.
	Personnel		1+		This group describes the primary contact person at the data centre who is able to respond to requests for and queries about the data. Most fields are self-explanatory. Address should be a valid postal address for the named person. Email, Phone and FAX fields are optional and may be repeated. These contact personnel include Investigator, Technical Contact, or DSD Author.
		Role	1+	Keyword	The Role may be repeated per personnel group. In other words, a person may have more than one role. In this case, the Personnel group needs only to be specified once in the DSD with repeating Role fields. Investigator: Person who headed the investigation or experiment that resulted in the acquisition of the data described (i.e., Principal Investigator, Experiment Team Leader) Technical Contact: A person who is knowledgeable about the technical content of the data (quality, processing methods, units, available software for further processing) DSD Author: The person who is responsible for the content of the DSD. If the responsibility shifts from the original author to another person, the DSD Author field should be updated to the new responsible person
		First_Name	1	Char(80)	Primary contact name
		Middle_Name	0|1	Char(80)	Primary contact name
		Last_Name	1	Char(80)	Primary contact name
		Email	1	Char(80)	Primary contact name
		Phone	1	Char(80)	Phone and fax numbers should be easily read by users. For example, for USA phone numbers, use either (xxx) xxx-xxx or xxx-xxx-xxxx. For other countries, begin the number with the country code: +44 123 45678
		Fax	0|1	Char(80)	
		Address	1	Char(512)	Primary contact address. The country should always be included in the address.

Field Name and attribute description	Cardinality	Type	Description
Reference	0|1	Char[1024]	Key bibliographic references pertaining to the data set described in the DSD. Bibliographic references may be provided in styles used by professional scientific journals. Hyperlinked URLs to online articles may be imbedded in the text by surrounding the URL in single or double quotes (http://earth.agu.org/eos_elec/95127e.html)
Summary	1	Char[6000]	A brief description of the data set, descriptive enough to allow potential users of the data set to determine if the data set is useful for their needs. Should include information needed for a user to determine the usefulness of the data set. Should start with a topic sentence, describing what information is in the data set. Often, this is some measurable quantity or quantities, such as sea surface temperature, human population density, or species mortality rate. The total length should consist of approximately 30 lines, with each line not exceeding 80 characters, and separated by a carriage return (or carriage return + line feed) character. Should restate information that may be found in other fields if that information is vital to the understanding of the data set (e.g., parameters, spatial coverage). Should reference the source information if the summary was abstracted from an existing document. Single spaced with blank lines separating paragraphs Capitalization should follow standard constructs. For readability, do not use all capital letters or all lower case letters, but use the appropriate case where applicable. No right justification Acronyms should be expanded to improve understanding. May contain tabular information Hyperlinked URLs may be embedded in the text by surrounding them with single or double quotes: " http://www.ngdc.noaa.gov " Where applicable, should include brief statements of the following important information: Data processing information (gridded, binned, swath, raw, algorithms used, necessary ancillary data sets) Methodology or analytical tools Time gaps in data set coverage Units and unit resolution Similarities and differences of these data to other closely-related data sets Other pertinent information Data format documentation (preferably as a URL)

Field Name and attribute description	Cardinality	Type	Description
Related_URL	0+		This field provides hypertext URL links to Internet sites that contain information related to the subject of the data, as well as other useful Internet sites such as project home pages, related data archives/servers (for instance direct manual FTP access to the directory containing the dataset, for those wishing to download the entire data set), metadata extensions, online software packages, and calibration/validation data.
URL_Content_Type	0|1	Keyword [Char[31]]	Describes the type of URL being referenced. This field is used only in the case where a database search of a specific URL type is necessary . URL_Content_Type should be selected from the URL Content Type Valids List (refer to : http://gcmd.gsfc.nasa.gov/Resources/valids/url_type.html . Most often, URL Content Type will not be entered. Rather, descriptive information about the associated data resource should be entered in the Description field.
URL	1	Char[256]	Hypertext link to the remote resource associated with the data set. This can repeat as required
Description	0|1	Char[1024]	Up to several lines of text describing the kind of information offered by the URL and associated with the data set.
DSD_File_Version	1	Int	The version of this metadata record. Version 1 is the original and DSD_File_Version number should be incremented for each new release of the metadata pointed to by this metadata record
DSD_Creation_Date	1	Char[17]	The date the metadata was created. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
DSD_Last_Revision_Date	1	Char[17]	The date the metadata was last revised. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
DSD_Revision_History	1	Char[1024]	A listing of changes made to the DSD over time. Provides a mechanism for tracking revisions to DSD content.

A6.3.1.5 MMR_DSD sample XML file

Before any MMR_FR can be ingested at the MMR system, a MMR_DSD must exist. MMR_DSD will be produced and maintained by RDAC and the GHRSST-PO in collaboration with the MMR system administrator. MMR_DSD will be generated as XML files according to the DTD provided in A6.3.1.2 An example of a GHRSST-PP MMR_DSD XML record is provided below.


```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE MMR_DSD SYSTEM "http://ghrsst-pp.metoffice.com/mmr_dsd.dtd">
<MMR_DSD>
  <Entry_ID>EUR-CMS-AVHRR16_L-IR-L2P-MOCC</Entry_ID>
  <Entry_Title>Sea Surface Temperature from AVHRR onboard NOAA-16, 2km resolution, over West Mediterranean sea</Entry_Title>
  <Data_Set_Citation>
    <Dataset_Creator>Meteo-France/CMS</Dataset_Creator>
    <Dataset_Title>Sea Surface Temperature from AVHRR onboard NOAA-16, 2km resolution, over West Mediterranean sea</Dataset_Title>
    <Dataset_Release_Date>20040120T120000Z</Dataset_Release_Date>
    <Dataset_Release_Place>Plouzane, France</Dataset_Release_Place>
    <Dataset_Publisher>IFREMER/CERSAT</Dataset_Publisher>
    <Version>1.0</Version>
    <Other_Citation_Details>These data are made available in the frame of the Medspiration project, funded by ESA</Other_Citation_Details>
    <Online_Resource>ftp://ftp.ifremer.fr/ifremer/cersat/medspiration/products/avhrr16_l/mocc/</Online_Resource>
  </Data_Set_Citation>
  <Parameters>
    <Category>Earth Science</Category>
    <Topic>Oceans</Topic>
    <Term>Ocean Temperature</Term>
    <Variable>Sea Surface Temperature</Variable>
    <Detailed_Variable>Skin Sea Surface Temperature</Detailed_Variable>
  </Parameters>
  <Sensor_Name>
    <Short_Name>AVHRR</Short_Name>
    <Long_Name>Advanced Very High Resolution Radiometer</Long_Name>
  </Sensor_Name>
  <Source_Name>
    <Short_Name>NOAA-16</Short_Name>
    <Long_Name>National Oceanic &amp; Atmospheric Administration-16</Long_Name>
  </Source_Name>
  <Temporal_Coverage>
    <Start_Date>20040913T120000Z</Start_Date>
  </Temporal_Coverage>
  <Spatial_Coverage>
    <Southernmost_Latitude>29.08</Southernmost_Latitude>
    <Northernmost_Latitude>47.51</Northernmost_Latitude>
    <Westernmost_Longitude>-8.32</Westernmost_Longitude>
    <Easternmost_Longitude>18.85</Easternmost_Longitude>
  </Spatial_Coverage>
  <Location>
    <Location_Name>Mediterranean Sea</Location_Name>
    <Detailed_Location>West Mediterranean Sea</Detailed_Location>
  </Location>
  <Projection_Information>
    <Projection_Type>Polar Stereographic</Projection_Type>
    <Ellipsoid_Type>International 1924</Ellipsoid_Type>
    <Other_Projection_Details>y axis is meridian 0</Other_Projection_Details>
  </Projection_Information>
  <Data_Resolution>
    <Latitude_Resolution>2 km</Latitude_Resolution>
    <Longitude_Resolution>2 km</Longitude_Resolution>
    <Temporal_Resolution>6 hours</Temporal_Resolution>
  </Data_Resolution>
  <Access_Constraints>Data available to users within 6 hours after acquisition</Access_Constraints>
  <Use_Constraints>Data may not be used for commercial applications.</Use_Constraints>
  <Originating_Center>METEO-FRANCE / Centre de Meteorologie Spatiale (CMS)</Originating_Center>
  <Data_Center>
    <Data_Center_Name>
      <Short_Name>IFREMER / CERSAT</Short_Name>
      <Long_Name>Institut Francais de Recherche pour l'Exploitation de la Mer / Centre ERS d'Archivage et de Traitement</Long_Name>
    </Data_Center_Name>
    <Personnel>
      <Role>Technical Contact</Role>
      <First_Name>Jean-Francois</First_Name>
```

```
            <Last_Name>Piolle</Last_Name>
            <Email>jfpiolle@ifremer.fr</Email>
            <Phone>555-333444</Phone>
            <Fax>555-333555</Fax>
            <Address>IFREMER CERSAT, Technopole Brest-Iroise, 29280 Plouzane, France </Address>
        </Personnel>
    </Data_Center>
    <Reference>Medspiration products user manual, Robinson I., Leborgne P., Piolle J.F., Larnicol G., v1.02, September 2004</Reference>
    <Summary>This dataset features remapped high resolution AVHRR data from NOAA-16 over west mediterranean</Summary>
    <Related_URL>
        <URL>http://www.medspiration.org</URL>
        <Description>This is the Home Page for the Medspiration Project. Additional project data, documentation, and software are available from this World Wide Web site.</Description>
    </Related_URL>
    <Related_URL>
        <URL_Content_Type>DODS Data Server</URL_Content_Type>
        <URL>http://www.ifremer.fr/cersat/dods/medspiration/</URL>
    </Related_URL>
    <Related_URL>
        <URL>ftp://ftp.ifremer.fr/ifremer/cersat/medspiration/</URL>
        <Description>The Medspiration FTP archive</Description>
    </Related_URL>
    <DSD_File_Version>1.0</DSD_File_Version>
    <DSD_Creation_Date>20040714T012234Z</DSD_Creation_Date>
    <DSD_Last_Revision_Date>20041023T123344Z</DSD_Last_Revision_Date>
    <DSD_Revision_History>
        2004-10-23, changed DODS url
        2004-09-26, changed technical contact phone number
        2004-09-15, updated dataset summary
    </DSD_Revision_History>
</MMR_DSD>
```


A6.3.2 MMR_FR metadata record format

A6.3.2.1 MMR_FR file name convention

MMR-FR files for L2P and L4/UHR products shall be named as:

FR-<filename>.xml

Where **<filename>** is the name of the related L2P or UHR/L4 data file without its format extension. For example:

FR-20030621-EUR-L4UHfnd-MED-v01.xml

;

A6.3.2.2 MMR_FR XML DTD

MMR-FR files shall be formatted in XML format, complying to the following DTD:

```
<!-- DTD for the GHRSST-PP MMR_FR                -->
<!--                                              -->
<!-- Author: J. F Piolle and C J Donlon           -->
<!--                                              -->
<!-- History                                      -->
<!-- 2004-02-17      Original                     -->
<!-- 2005-01-11      Revised time  formats and character lengths        -->


<!ELEMENT MMR_FR ( Entry_ID,
                File_Name,
                File_Release_Date,
                File_Version,
                Related_URL+,
                Temporal_Coverage,
                Spatial_Coverage,
                Personnel+,
                        Metadata_History
 ) >
```

```
<!-- DSD_Entry_ID (reference) ===================================== -->
<!ELEMENT Entry_ID (#PCDATA)>


<!ELEMENT File_Name  (#PCDATA)>
<!ELEMENT File_Release_Date (#PCDATA)>
<!ELEMENT File_Version (#PCDATA)>

<!-- Related URL ================================================== -->
<!ELEMENT Related_URL ( URL_Content_Type?,
                    URL+,
                    Description?
                   )
>

<!ELEMENT URL_Content_Type (#PCDATA)>
<!ELEMENT URL (#PCDATA)>
<!ELEMENT Description (#PCDATA)>

<!-- Temporal Coverage ============================================ -->
<!ELEMENT Temporal_Coverage ( Start_Date,
                          Stop_Date
                         )
>

<!ELEMENT Start_Date (#PCDATA)>
<!ELEMENT Stop_Date (#PCDATA)>


<!-- Spatial Coverage ============================================= -->
<!ELEMENT Spatial_Coverage ( Southernmost_Latitude,
                         Northernmost_Latitude,
                         Westernmost_Longitude,
                         Easternmost_Longitude,
                         Minimum_Altitude?,
                         Maximum_Altitude?,
                         Minimum_Depth?,
                         Maximum_Depth?
                        )
>

<!ELEMENT Southernmost_Latitude (#PCDATA) >
<!ELEMENT Northernmost_Latitude (#PCDATA) >
<!ELEMENT Westernmost_Longitude (#PCDATA) >
<!ELEMENT Easternmost_Longitude (#PCDATA) >
<!ELEMENT Minimum_Altitude (#PCDATA) >
<!ELEMENT Maximum_Altitude (#PCDATA) >
<!ELEMENT Minimum_Depth (#PCDATA) >
<!ELEMENT Maximum_Depth (#PCDATA)>

<!-- Personnel ==================================================== -->
<!ELEMENT Personnel ( Role+,
                  First_Name,
                  Middle_Name?,
                  Last_Name,
                  Email,
                  Phone,
                  Fax?,
                  Address
                 )
>

<!ELEMENT Role (#PCDATA)>
<!ELEMENT First_Name (#PCDATA)>
<!ELEMENT Middle_Name (#PCDATA)>
<!ELEMENT Last_Name (#PCDATA)>
<!ELEMENT Email (#PCDATA)>
<!ELEMENT Phone (#PCDATA)>
<!ELEMENT Fax (#PCDATA)>
<!ELEMENT Address (#PCDATA)>

<!-- Metadata_History ============================================= -->
<!ELEMENT Metadata_History (FR_File_Version,
                           FR_Creation_Date,
                           FR_Revision_History,
                           FR_Last_Revision_Date)>
```

```
<!--  Creation Date ============================================ -->
<!ELEMENT FR_File_Version (#PCDATA)>

<!-- Creation Date ============================================ -->
<!ELEMENT FR_Creation_Date (#PCDATA)>

<!-- Last Revision Date ======================================== -->
<!ELEMENT FR_Last_Revision_Date (#PCDATA)>

<!--  Revision History ========================================= -->
<!ELEMENT FR_Revision_History (#PCDATA)>
```

A6.3.2.3 MMR_FR DTD description

Table A6.3.2.3 provides a more explicit and readable description of the DTD specified in A6.3.2.2.

Table 6.3.2.3 XML DTD for MMR_FR metadata records

Field Name and attribute description	Cardinality	Type	Description
Entry_ID	1	Char(160)	Identical to Entry_ID field defined in Table A6.3.1
File_Name	1	Char(160)	Data provider filename pointed to by this metadata record
File_Release_Date	1	Char[17]	The date the data file pointed to by this metadata record was created. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
File_Version	1	Int	The version of data file pointed to by this metadata record of this metadata record.
Related_URL	0+	group	This field provides hypertext URL links to Internet sites that contain information related to the subject of the data, as well as other useful Internet sites such as project home pages, related data archives/servers (for instance direct manual FTP access to the directory containing the dataset, for those wishing to download the entire data set), metadata extensions, online software packages, and calibration/validation data.
URL_Content_type	0 | 1	Keyword [Char[31]]	Describes the type of URL being referenced. This field is used only in the case where a database search of a specific URL type is necessary . URL_Content_Type should be selected from the URL Content Type Valids List (refer to : http://gcmd.gsfc.nasa.gov/Resources/valids/url_type.html . Most often, URL Content Type will not be entered. Rather, descriptive information about the associated data resource should be entered in the Description field.
URL	1+	Char[256]	Hypertext link to the remote resource associated with the data set. This can repeat as required
Description	0 | 1	Char[1024]	Up to several lines of text describing the kind of information offered by the URL and associated with the data set.
Temporal_Coverage	1	group	*This group* provides the start and end date and times for the measurements contained in the GHRSST-PP data file. All date and times must be in Universal Time Coordinates (UTC, which is approx. Greenwich Mean Time), and formatted at "yyyy-mm-dd" for dates and "hh:mm:ss" for times.

Field Name and attribute description	Cardinality	Type	Description
Start_Date	1	Char[17]	ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
Stop_Date	1	Char[17]	ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
Spatial_Coverage	1	group	Geographic coverage (horizontal and vertical) of the data described. Longitude must be expressed in the range -180.0 to +360.0 as an offset from the Greenwich meridian, and latitude in the range -90.0 to +90.0, as an offset from the Equator. For gridded data, these coordinates should be the outer edges of the coverage, and not the coordinates of the centres of the edge pixels.
Southernmost_Latitude	1	Float	DD.ddd
Northernmost_Latitude	1	Float	DD.ddd
Westernmost_Longitude	1	Float	DDD.ddd
Easternmost_Longitude	1	Float	DDD.ddd
Minumum_Altitude	0 | 1	Char[80]	The altitude level specified in Meters which represents the lower limit of data coverage, as measured from mean sea level (Primarily for Model outputs)
Maximum_Altitude	0 | 1	Char[80]	
Minumum_Depth	0 | 1	Char[80]	The depth level specified in Meters which represents the upper-most depth of data coverage, as measured from mean sea level. (Primarily for Model outputs)
Maximum_Depth	0 | 1	Char[80]	
Personnel	1+		This group describes the primary contact person at the data centre who is able to respond to requests for and queries about the data. Most fields are self-explanatory. Address should be a valid postal address for the named person. Email, Phone and FAX fields are optional and may be repeated. These contact personnel include Investigator, Technical Contact, or DSD Author.
Role	1+	Keyword	The Role may be repeated per personnel group. In other words, a person may have more than one role. In this case, the Personnel group needs only to be

Field Name and attribute description	Cardinality	Type	Description
			specified once in the DSD with repeating Role fields. • **Investigator**: Person who headed the investigation or experiment that resulted in the acquisition of the data described (i.e., Principal Investigator, Experiment Team Leader) • **Technical Contact**: A person who is knowledgeable about the technical content of the data (quality, processing methods, units, available software for further processing) • **DSD Author**: The person who is responsible for the content of the DSD. If the responsibility shifts from the original author to another person, the DSD Author field should be updated to the new responsible person
First_Name	1	Char(80)	Primary contact name
Middle_Name	0 | 1	Char(80)	Primary contact name
Last_Name	1	Char(80)	Primary contact name
Email	1	Char(80)	Primary contact name
Phone	1	Char(80)	Phone and fax numbers should be easily read by users. For example, for USA phone numbers, use either (xxx) xxx-xxx or xxx-xxx-xxxx. For other countries, begin the number with the country code: +44 123 45678
Fax	0 | 1	Char(80)	
Address	1	Char(512)	Primary contact address. The country should always be included in the address.
Metadata_History	1	group	
FR_File_Version	1	Int	The version of this metadata record. Version 1 is the original and File_Version number should be incremented for each new release of the metadata. Version 1 is the original and FR_File_Version number should be incremented for each new release of the metadata. This value will determine if a MMR record is updated
FR_Creation_Date	1	Char[10]	The date the metadata was created. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
FR_Last_Revision_Date	1	Char[17]	The date the metadata was last revised. ISO-8601 standard date/time string in the format: yyyymmddThhmmssZ, e.g., 20050111T123344Z where yyyy=year mm = month

Field Name and attribute description		Cardinality	Type	Description
				dd = day of month T = separator between date and time that should follow ther date hh = hour of day (0-23) mm = minute of the hour (0-59) ss = second of the minute (0-59) Z = time zone identifier for 00 UTC
	FR_Revision_History	1	Char[1024]	A listing of changes made to the DSD over time. Provides a mechanism for tracking revisions to DSD content.

A6.3.2.4 MMR-FR sample XML file

Each MMR_FR metadata record should be encoded in extensible mark-up language (XML) ASCII text format. An example MMR_FR metadata record formatted as an XML document according to the DTD provided in A6.3.2.2 is given below.


```xml
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE MMR_FR SYSTEM "http://ghrsst-pp.metoffice.com/mmr_fr.dtd">
<MMR_FR>
  <Entry_ID>EUR-CMS-AVHRR16_L-IR-L2P-MOCC</Entry_ID>
  <File_Name>20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-LAC20030621A7SST-v01.nc</File_Name>
  <File_Release_Date>20040913T123344Z</File_Release_Date>
  <File_Version>1.0</File_Version>
  <Related_URL>
    <URL_Content_Type>DODS Data Server</URL_Content_Type>
    <URL>http://www.ifremer.fr/cersat/dods/medspiration/20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-
LAC20030621A7SST-v01.nc</URL>
  </Related_URL>
  <Related_URL>
    <URL>ftp://ftp.ifremer.fr/ifremer/cersat/medspiration/20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-
LAC20030621A7SST-v01.zip</URL>
    <Description>The Medspiration FTP archive</Description>
  </Related_URL>
  <Temporal_Coverage>
    <Start_Date>20040913T103315Z</Start_Date>
    <Stop_Date>20040913T104522Z</Stop_Date>
  </Temporal_Coverage>
  <Spatial_Coverage>
    <Southernmost_Latitude>29.080</Southernmost_Latitude>
    <Northernmost_Latitude>47.510</Northernmost_Latitude>
    <Westernmost_Longitude>-8.320</Westernmost_Longitude>
    <Easternmost_Longitude>18.850</Easternmost_Longitude>
  </Spatial_Coverage>
  <Personnel>
    <Role>Technical Contact</Role>
    <First_Name>Jean-Francois</First_Name>
    <Last_Name>Piolle</Last_Name>
    <Email>jfpiolle@ifremer.fr</Email>
    <Phone>555-333444</Phone>
    <Fax>555-333555</Fax>
    <Address>IFREMER CERSAT, Technopole Brest-Iroise, 29280 Plouzane, France</Address>
  </Personnel>
 <Metadata_History>
    <FR_File_Version>1.0</FR_File_Version>
    <FR_Creation_Date>20041023T123344Z</FR_Creation_Date>
    <FR_Last_Revision_Date>20041023T123344Z</FR_Last_Revision_History>
    <FR_Revision_History>
      2004-10-23, first file transfer
    </FR_Revision_History>
 </Metadata_History>
</MMR_FR>
```


A6.4 Registering MMR_FR metadata record at the MMR via e-mail

GDS MMR_FR are registered with the GHRSST-PP MMR system via e-mail. Each XML encoded MMR_FR is entered as the body of an ASCII text e-mail message which is then sent to:

ghrsst_mmr@podaac.jpl.gov

The subject line of each metadata message should be formatted as follows:

```
Subject: GHRSST metadata notification from <processing centre code>
```

where <processing centre code> is defined in Table 2.7.1. For example

```
Subject: GHRSST metadata notification from JAP
```

would indicate that the e-mail contains a metadata record from the Japanese RDAC system. Figure 6.4.1 shows an example metadata delivery message header.

```
Subject: GHRSST metadata notification from EUR
Date: Thu, 30 Jan 2003 17:42:17 -0800
From: Jorge Vazquez <jv@pacific.jpl.nasa.gov>
Organization: JPL/Caltech
To: GHRSST-PP MMR <MMR_ghrsst@podaac.jpl.gov>
References: <3E385EC5.18CF9B3F@seanet.jpl.nasa.gov>
```

Figure A6.4.1 An example e-mail header for a MMR_FR delivery message.

If the MMR XML parser rejects the MMR_FR metadata message, it will return an e-mail with a full diagnostic report explaining the cause of the error to the return address specified in the delivery e-mail. A revised metadata message should be prepared and sent to the MMR as soon as possible.

On successful parsing of the MMR data record, the GHRSST-PP MMR system will send a confirmation e-mail to the

A6. 5 Modifying GDS metadata records already registered at the MMR

If an incorrect MMR_FR metadata record is sent to the MMR or an MMR_FR needs to be revised, a new metadata message is sent to the MMR. The MMR_FR element **`File_Version`** should be incremented to indicate that this metadata record should replace any existing data already registered at the MMR under this **`Entry_ID`**. **CAUTION: This action will replace any previous version of the metadata record held in the MMR database system with the revised version.**

A6.6 Retrieval of metadata from the MMR

MMR data may be retrieved via a web browser interface.

Appendix A7. GDS operation and error messages

The GDS operational log is a central e-mail managed system that collates and publishes significant operations within the GDS. Each time a relevant error/warning is generated at an RDAC or GDAC processor, an e-mail is sent to the GDS error log errlog@ghrsst-pp.org and the message contents automatically published on the GHRSST-PO ERRLOG web page located at http://www.ghrsst-pp.org/ERRLOG. Future versions of the GDS may use alternative methods to deliver error messages (e.g., ftp).

All operation and error message are formatted as xml data files according to the specifications described in the following sections.

A7.1 Operation and error message file naming convention

The operation and error message files shall be sent with the following name (or subject if sent by email) :

<Processing Centre Code>-<Message type>-<time stamp>-<Code>.<base format>

Table A7.1.1 GDS errlog filenaming convention.

Name	Definition	Description
<Message type>	ERR=error REQ=request WAR=warning OPE=operation DEB=debug	Type of message
<Processing Centre Code>	Refer to GDS Appendix A2 Table A2.1	Processing centre code
<time stamp>	yyyymmddThhmmssZ	Date of creation of message UTC
<code>	Refer to message tables	The code of the message
<base format>	Xml	Generic file format

e.g., **`EUR-ERR-2004042T5022356Z.124.xml`** as defined in Table A7.1.1.

A7.2 Operational and error message XML DTD

The operation and error messages exchanged between the various systems and the controller within Medspiration and between Medspiration and the GDAC shall be formatted in XML format, complying to the following DTD where the body element is intended for all informational context related to a message and will be refined as protocol specifications progress.

```
<!-- DTD for Medspiration messages  -->

<!ELEMENT op_message (sender, recipient, date, reference?, body?)>

<!-- type of message -->
<!ATTLIST op_message type (error|request|acknowledgment|warning|operation) "error">

<!-- code of message (refer to the tables below) -->
<!ATTLIST op_message code CDATA #REQUIRED>

<!-- Entry sender ========================================== -->
<!ELEMENT sender (#PCDATA)>
<!-- type of sender : auto (process), operator (sent manually by a technician), unknown -->
<!ATTLIST sender type (auto | operator | unknown) #REQUIRED>


<!-- Entry recipient ========================================== -->
<!ELEMENT recipient (#PCDATA)>
```

```
<!-- type of recipient : auto (process), operator (requires manual analysis by a technician), unknown -->
<!ATTLIST recipient type (auto | operator | unknown) #REQUIRED>

<!-- Entry date  ============================================ -->
<!ELEMENT date EMPTY>
<!ATTLIST date year CDATA  #REQUIRED>
<!ATTLIST date month CDATA  #REQUIRED>
<!ATTLIST date day CDATA  #REQUIRED>
<!ATTLIST date hour CDATA  #REQUIRED>
<!ATTLIST date minute CDATA  #REQUIRED>
<!ATTLIST date second CDATA  #REQUIRED>
<!ATTLIST date millisec CDATA  #REQUIRED>
<!ATTLIST date timeZone (UTC) #IMPLIED>

<!-- Entry reference ======================================== -->
<!ELEMENT reference (#PCDATA)>

<!-- type of reference : data (file) or message (another message) -->
<!ATTLIST reference type (data | message) #REQUIRED>


<!-- Entry body =========================================== -->
<!ELEMENT body ((missing_field+ | error_detail)?, comment?)>

<!ELEMENT missing_field (#PCDATA)>
<!ELEMENT error_detail (#PCDATA)>
<!ELEMENT comment (#PCDATA)>
```

A7.3 Operational and error message XML DTD description

Table A7.3.1 provides a more explicit and readable description of the operational and error message XML DTD. It is fully consistent with the DTD described in section A7.2.

Table A7.3.1 Operational and error message XML DTD description

Table Colour code key	
(blue)	Attribute
(orange)	Element
(green)	Denotes elements that must not be present together (case choice)

Field Name and attribute description				Cardinality	Type	Description
op_message				*1*	*group*	*'op_message' is the envelop of the message. The 'type' and 'code' attributes must be specified :* *'type' is the category of the message* *'code' is the identifier of the message refering to message tables defined for each type of message (error, request,…)*
type	*Req*	*string*	*error* *request* *warning* *acknowledgment* *operation*			
code	*Req*	*int*	Refer to message tables			
sender				*1*	*string*	*'sender' is the system or centre (GDAC, RDACs) which sent the message. The 'type' attribute must be specified. The values for this subgroup must be agreed by the two communicating entities in order to ensure suitable analysis and processing of the message.* *The 'type' of sender can be 'auto' (automatic process), 'operator' (human technician in charge of the system) or unknown.*
type	*Req*	*string*	*auto* *operator* *unknown*			
recipient				*1*	*string*	*'recipient is the system or centre (GDAC, RDACs) for which the message is intended. The 'type' attribute must be specified. The values for this subgroup must be agreed by the two communicating entities in order to ensure suitable analysis and processing of the message.* *The 'type' of recipient targeted can be 'auto' (automatic process), 'operator' (human technician in charge of the recipient system) or unknown.*
type	*Req*	*string*	*auto* *operator* *unknown*			

	date *year* | *Req* | *int* *month* | *Req* | *int* *day* | *Req* | *int* *hour* | *Req* | *int* *minute* | *Req* | *int* *second* | *Req* | *int* *millisec* | *Req* | *int* *timeZone* | *Req* | *string* | *UTC*	*1*	*empty*	*The date of creation of the message. All date information are self-contained into the attributes.*
	reference *type* | *Req* | *string* | *data message*	*0|1*	*string*	*The 'reference' element contains a reference to a related data file (for instance when requesting the retransmission of a data file) or message (for instance when replying to a previous message). The type of referenced file is set in the 'type' attribute. when refering to a data file, the value shall be the name of the data file. When refering to a message, the value shall be the name (or subject) of the message.*
	body	*0|1*	*subgroup*	*This subgroup is intended for a complete description of the purpose of the message if needed.* *The 'missing_field' or 'error_detail' are two alternative subelements related each to a specific type of message (respectively warning and error). They shall be used for these messages only and never together (refer to the message tables).*
	missing_field	*1+*	*string*	*The name of the missing field when sending an uncomplete data file. The element shall be repeated when several fields are missing.*
	error_detail	*1*	*string*	*An error string*
	To be defined			To be defined
	comment	*0|1*	*string*	*A free text*

A7.4 Format of GDS operations and error messages

RDAC/GDAC error message files shall be specified accordfing to Table A7.4.1. The body content is likely to evolve as the list of potential error messages increases.

Table A7.4.1 GDS Error message format specifications

Code	Description	Reference	Body content
0	Free format user specified	None	`<!ELEMENT body (comment)> <!ELEMENT comment(#PCDATA)>`
1	RDAC/GDAC processing system is down	None	`<!ELEMENT body (error_detail)> <!ELEMENT error_detail (#PCDATA)>` *More to be defined as required*
2	RDAC/GDAC processing system is back to operation	None	
3	L2 rejected	L2 filename	
4	L2P generation failed	L2P filename	
5	L4 generation failed	L4 filename	
6	MMR-FR rejected	MMR-FR filename	
7	MMR-DSD rejected	MMR-DSD filename	
8	MDB record rejected	MDB record filename	
9	Transmission failed	Filename	
10	File not correctly formatted	Filename	
11	Provider server is down	DSD name	
12	Provider directory not accessible	DSD name	
13	L2P rejected	L2P filename	
14	Ancillary data rejected	Filename	
13	*To be defined*	*To be defined*	

GDS system log acknowledgement message files shall be specified according to Table A7.4.2.

Table A7.4.2 GDS System acknowledgement message format specifications

code	Description	Reference	Body content
0	Free format user specified	free	`<!ELEMENT body (comment)> <!ELEMENT comment(#PCDATA)>`
1	L2P successfully transmitted	L2P filename	No body content
2	Ancillary data successfully transmitted	Ancillary data filename	
3	*To be defined*	*To be defined*	

A request message (e.g., a Data Ready Notification (DRN) message) file shall be specified according to Table A7.4.3.

Table A7.4.3 GDS request message format specification.

code	Description	Reference	content
0	Free format user	None	`<!ELEMENT body (#PCDATA)>`

	specified		<!ELEMENT comment(#PCDATA)>
1	L2P retransmission requested	L2P filename	No body content
2	Ancillary data retransmission requested	Ancillary data filename	
3	*To be defined*	*To be defined*	

GDS warning message files shall be specified according to Table A7.4.4.

Table A7.4.4 GDS warning message format specifications

code	Description	Reference	content
0	free	free	<!ELEMENT body (comment)> <!ELEMENT comment(#PCDATA)>
1	L2P product not complete	L2P filename	<!ELEMENT body (missing_field+)> <!ELEMENT missing_field (wind_speed | …)> *To be completed*
2	L4 product not complete	L4 filename	
3	product out of delay	L2/L2P/ancillary filename	*To be defined*
4	Dataset problematic (no file of the corresponding dataset has arrived before cut-off)	L2 or ancillary data filename	*To be defined*
5	*To be defined*	*To be defined*	*To be defined*

GDS operations message files shall be specified according to Table A7.4.6.

Table A7.4.6 GDS warning message format specifications

code	Description	Reference	content
0	free	free	<!ELEMENT body (comment)> <!ELEMENT comment(#PCDATA)>
1	*To be defined*	*To be defined*	*To be defined*

A7.5 Sample files (XML) for GDS operations and error messages

A7.5.1 Error message

This error message warns the L2P file was rejected (rain contamination).

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE op_message SYSTEM "op_message.dtd">
<op_message type="error" code="13">
   <sender type="auto">EUR</sender>
   <recipient type="auto">GDAC</recipient>
   <date year="2004" month="02" day="25"
        hour="16" minute="00" second="58" millisec="156"/>
   <reference type="data">20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-LAC20030621A7SST-v01.nc
   </reference>
   <body>
      <error_detail>rain rejection rate critical</error_detail>
   </body>
</op_message>
```

A7.5.2 Acknowledgement message

This message is an acknowledgement for a L2P data file correct transmission.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE op_message SYSTEM "op_message.dtd">
```

```
<op_message type="acknowledgment" code="1">
   <sender type="auto">IFREMER</sender>
   <recipient type="auto">CMS</recipient>
   <date year="2004" month="02" day="25"
        hour="16" minute="00" second="58" millisec="156"/>
   <reference                  type="data">20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-LAC20030621A7SST-v01.nc
   </reference>
</op_message>
```

A7.5.3 Request message

This message is a request for a L2P data file retransmission.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE op_message SYSTEM "op_message.dtd">

<op_message type="request" code="1">
   <sender type="auto">IFREMER</sender>
   <recipient type="auto">CMS</recipient>
   <date year="2004" month="02" day="25"
        hour="16" minute="00" second="58" millisec="156"/>
   <reference                  type="data">20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-LAC20030621A7SST-v01.nc
   </reference>
</op_message>
```

A7.5.4 Warning message

This message is a warning stating a L2P data file is provided uncomplete (missing variables).

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE op_message SYSTEM "op_message.dtd">
<op_message type="warning" code="1">
   <sender type="auto">EUR</sender>
   <recipient type="auto">GDAC</recipient>
   <date year="2004" month="02" day="25"
        hour="16" minute="00" second="58" millisec="156"/>
   <reference                  type="data">20030621-EUR-CMS-AVHRR16_L-IR-L2P-MOCC-LAC20030621A7SST-v01.nc
   </reference>
   <body>
      <missing_field>aerosol_optical_depth</missing_field>
      <missing_field>surface_solar_irradiance</missing_field>
   </body>
</op_message>
```

Appendix A8 The GDS Problem resolution board

A Problem Resolution Board (PRB) is based on the ARM approach. There may be other, possibly better, ways of solving GDS problems but this approach does provide traceability of decisions, and helps avoid someone making a change to solve one problem but in so doing unintentionally creating another one through being unaware of downstream implications.

The problem resolution board should contain the following sub-groups:

- **Cal-Val problem resolution board**
- **Data management problem resolution board**
- **GDS problem resolution board**
- **ISDI-OPS and ERRLOG log monitors**

The PRB must establish for each problem situation:

- What is the control loop to decide on what to do?
- How should the ISDI-TAG deal with this?
- What are the allowable timescales?
- How should operational users be informed of the problem and how quickly can they be informed and by which route (e-mail??)
- What is the process for rectifying anomalies?
- What is the control loop to decide on what to do?
- How should the ISDI-TAG deal with this?
- What are the allowable timescales?
- How should operational users be informed of the problem and how quickly can they be informed and by which route (e-mail??)

The GHRSST-PP GDS PRB will be convened by the GHRSST-PP International Project Office once the GDS is operational. In addition, a PRB may be convened at any time as required to solve specific problems.

Appendix A9 The GDS System Configuration file

The GDS system configuration file contains a set of system wide configuration values that are expected to evolve throughout the lifetime of the GDS. Each RDAC will require a copy of the configuration file in order to maintain consistency across the distributed processing system. The GHRSST-PP International Project Office (GHRSST-PO) will maintain the master copy of the GDS configuration file which will be made available on the GHRSST-PP web site.

A9.1 Configuration file naming convention

The following filename convention will be used for all versions of the GDS configuration file:

`GDS_V<version number>_R<revision number>_<release date>_<config file type>.config`

which is described in Table A9.1.

Table A9.1 Configuration filename components

Name	Definition	Description
`<version number>`	1.0	indicates the applicable GDS version number (e.g., 1.0
`<revision number>`	1.4	indicates the revision number for the particular version of the GDS (e.g., 1.4)
`<release date>`	yyyymmddThhmmssZ yyyy=year mm=month (1-12) dd=day of month (1-31) T=time string identifier hh=hour of the day (0-23) mm=minute of the hour (0-59) ss=second of the minute (0-59) Z=indicates time specification is UTC	Configuration file creation date
`<config file type>`	Table A9.2	Type co configuration file

Table A9.2 Code value indicating the type of GDS configuration file required to name a GDS configuration filename.

Code	Description
`qc_configuration`	Configuration file for the L2 and ancillary data qualty check
`ipcv_l2p_parameters`	Configuration file for the Infra-red Proximity confidence value derivation
`mpcv_l2p_parameters`	Configuration file for the Microwave Proximity confidence value derivation
`sses_parameters`	Configuration file for Single Sensor Error Statistics
`reference_parameters`	Configuration file for the reference parameters
`MDB_parameters`	Matchup database configuration parameters

For example

`GDS_V1.0_R1.4_20040225T122334Z_sses_parameters.config`

Would refer to a GDS SSES configuration file valid for GDS version 1, release 1.4 that was issued at 12:23:34 UTC on the 25th February 2004.

A9.2 L2 and ancillary data quality control configuration file

The configuration file for the L2 and ancillary data quality check shall be formatted in XML format, complying to the DTD provided in Figure

A9.2.1 L2 and ancillary data quality control configuration file XML DTD

```
<!-- DTD for Medspiration L2 Parameter Configuration file ================= -->

<!-- Entry qc_configuration ===================================================== -->
<!ELEMENT qc_configuration (limit*, max_rejection_rate*)>

<!-- Entry limit ======================================================= -->
<!ELEMENT limit (low?, high?)>
<!ATTLIST limit parameter (sst | wspd| aod | ssi) #REQUIRED>

<!-- Entry low ============================================================ -->
<!ELEMENT low (#PCDATA)>

<!-- Entry high =========================================================== -->
<!ELEMENT high (#PCDATA)>

<!-- Entry max_rejection_rate ============================================== -->
<!ELEMENT max_rejection_rate (#PCDATA)>
<!ATTLIST max_rejection_rate source (l2 | ancillary) #REQUIRED>
<!ATTLIST max_rejection_rate parameter
(sst | cloud | cosmetic | rain | ice | land | wspd | ssi | aod) #REQUIRED>
```

A9.2.2 L2 and ancillary data quality control configuration file XML sample file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE qc_configuration SYSTEM "qc_configuration.dtd">

<qc_configuration>
   <limit parameter="sst">
      <low>-3</low>
      <high>40</high>
   </limit>
   <limit parameter="wspd">
      <high>50</high>
   </limit>
   <limit parameter="ssi">
      <high>350</high>
   </limit>
   <limit parameter="aod">
      <low>10</low>
      <high>400</high>
   </limit>
   <max_rejection_rate source="l2" parameter="sst">80</max_rejection_rate>
   <max_rejection_rate source="l2" parameter="cloud">80</max_rejection_rate>
   <max_rejection_rate source="l2" parameter="cosmetic">80</max_rejection_rate>
   <max_rejection_rate source="l2" parameter="rain">80/max_rejection_rate>
   <max_rejection_rate source="l2" parameter="ice">80/max_rejection_rate>
   <max_rejection_rate source="l2" parameter="land">80/max_rejection_rate>
   <max_rejection_rate source="ancillary" parameter="wspd">80</max_rejection_rate>
   <max_rejection_rate source="ancillary" parameter="ssi">80</max_rejection_rate>
   <max_rejection_rate source="ancillary" parameter="aod">80</max_rejection_rate>
</qc_configuration>
```

A9.3 IPCV and L2P evaluation parameter configuration file

The following formats are specifieds for the IPCV and L2P evaluation parameters configuration file.

A9.3.1 IPCV and L2P evaluation parameter configuration file XML DTD

The configuration file for the Infra-red proximity confidence value derivation shall be formatted in XML format, complying to the following DTD.

```
<!-- DTD for Medspiration L2P Parameter Configuration file ================= -->

<!ELEMENT ipcv_l2p_parameters (dataset+)>


<!-- Entry dataset ============================================ -->
<!ELEMENT dataset (distance, threshold, max_erron_rate)>

<!-- Unique Id for each dataset ie dsd entry  =========================== -->
<!ATTLIST dataset l2_entry ID #REQUIRED>

<!ELEMENT distance (#PCDATA)>
<!ATTLIST distance column CDATA #REQUIRED>

<!ELEMENT threshold(#PCDATA)>
<!ATTLIST threshold row CDATA #REQUIRED>

<!ELEMENT max_erron_rate (#PCDATA)>
```

A9.3.2 IPCV and L2P evaluation parameter configuration file XML sample file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ipcv_l2p_parameters SYSTEM "ipcv_l2p_parameters.dtd">

<ipcv_l2p_parameters>
   <dataset l2_entry="AVHRR16_L">
      <threshold row=1>1</threshold >
      <threshold row=2>2</threshold >
      <threshold row=3>5</threshold >
      <distance column=1>100</distance >
      <distance column=2>200</distance >
      <max_erron_rate>3000</max_erron_rate>
   </dataset>
   <dataset l2_entry="AVHRR17_L">
      <threshold row=1>1</threshold >
      <threshold row=2>2</threshold >
      <threshold row=3>5</threshold >
      <distance column=1>100</distance >
      <distance column=2>200</distance >
      <max_erron_rate>3000</max_erron_rate>
   </dataset>
</ipcv_l2p_parameters>
```

A9.4 MPCVand L2P evaluation parameter configuration file

The following formats are specifieds for the MPCV and L2P evaluation parameters configuration file.

A9.4.1 MPCVand L2P evaluation parameter configuration file XML DTD

The configuration file for the Microwave Proximity confidence value derivation shall be formatted in XML format, complying to the following DTD.

```
<!-- DTD for Medspiration L2P Parameter Configuration file ================= -->

<!ELEMENT mpcv_l2p_parameters (dataset+)>


<!-- Entry dataset ============================================ -->
<!ELEMENT dataset (sidelobe_distance_threshold,
                relaxed_mw_rain_distance,
                relaxed_rain_threshold,
```

```
                wind_threshold,
                proximity)>

<!-- Unique Id for each dataset ie dsd entry  ============================ -->
<!ATTLIST dataset l2_entry ID #REQUIRED>

<!ELEMENT sidelobe_distance_threshold (#PCDATA)>
<!ELEMENT relaxed_mw_rain_distance (#PCDATA)>
<!ELEMENT relaxed_rain_threshold (#PCDATA)>

<!ELEMENT wind_threshold (#PCDATA)>
<!ATTLIST wind_threshold row CDATA #REQUIRED>

<!ELEMENT proximity (#PCDATA)>
```

A9.4.2 MPCVand L2P evaluation parameter configuration file XML sample file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE mpcv_l2p_parameters SYSTEM "mpcv_ l2p_parameters.dtd">

<mpcv_parameters>
   <dataset l2_entry="amsre">
      <sidelobe_distance_threshold>50</sidelobe_distance_threshold>
      <relaxed_mw_rain_distance>50</relaxed_mw_rain_distance>
      <relaxed_rain_threshold>50</relaxed_rain_threshold>
      <wind_threshold row=1>6</wind_threshold >
      <wind_threshold row=2>20</wind_threshold >
      <proximity></proximity>
   </dataset>
   <dataset l2_entry="tmi">
      <sidelobe_distance_threshold>50</sidelobe_distance_threshold>
      <relaxed_mw_rain_distance>50</relaxed_mw_rain_distance>
      <relaxed_rain_threshold>50</relaxed_rain_threshold>
      <wind_threshold row=1>6</wind_threshold >
      <wind_threshold row=2>20</wind_threshold >
   </dataset>
</mpcv_parameters>
```

A9.5 SSES configuration file

The following formats are specifieds for the SSES configuration file.

A9.5.1 SSES XML DTD

The configuration file for SSES shall be formatted in XML format, complying to the following DTD.

```
<!-- DTD for Medspiration L2P Parameter Configuration file ================= -->

<!ELEMENT sses_parameters (dataset+)>


<!-- Entry dataset ============================================= -->
<!ELEMENT dataset (sses+)>

<!-- Unique Id for each data set ie DSD name  ============================ -->
<!ATTLIST dataset dsd_entry ID #REQUIRED>

<!ELEMENT sses (bias, stdev)>
<!ATTLIST sses confidence_level CDATA #REQUIRED>

<!ELEMENT bias (#PCDATA)>
<!ELEMENT stdev (#PCDATA)>
```

A9.5.2 SSES XML sample file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE sses_parameters SYSTEM "sses_parameters.dtd">
<sses_parameters>
   <dataset dsd_entry="EUR-CMS-AVHRR16_L">
```

```
    <sses confidence_level="1">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="3">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="4">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="5">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="6">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
  </dataset>
  <dataset dsd_entry="EUR-CMS-AVHRR17_L">
    <sses confidence_level="1">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="3">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="4">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="5">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
    <sses confidence_level="6">
      <bias>0.12</bias>
      <stdev>0.34</stdev>
    </sses>
  </dataset>
</sses_parameters>
```


A9.6 Reference parameters configuration file

The following formats are specifieds for the reference parameters configuration file.

A9.6.1 Reference parameters configuration file XML DTD

The configuration file for the reference parameters shall be formatted in XML format, complying to the following DTD.

```
<!-- DTD for Medspiration L2P Parameter Configuration file ================= -->

<!ELEMENT reference_parameters (contamination, upwelling)>


<!-- Entry contamination ============================================ -->
<!ELEMENT contamination (sidelobe_distance_threshold)>
<!ELEMENT sidelobe_distance_threshold (#PCDATA)>

<!-- Entry upwelling ============================================ -->
<!ELEMENT upwelling (distance_to_coast+)>

<!ELEMENT distance_to_coast  (#PCDATA)>
```

```
<!ATTLIST distance_to_coast area (atlantic | mediterranean) #REQUIRED>
```

A9.6.2 Reference parameters configuration file XML sample file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE reference_parameters SYSTEM "reference_parameters.dtd">
<reference_parameters>
   <contamination>
      <sidelobe_distance_threshold>20</sidelobe_distance_threshold>
   </contamination>
   <upwelling>
      <distance_to_coast area="atlantic">100</distance_to_coast>
      <distance_to_coast area="mediterranean">50</distance_to_coast>
   </upwelling>
</reference_parameters>
```

A9.7 MDB parameters configuration file

The following formats are specifieds for the MDB parameters configuration file.

A9.7.1 MDB parameters configuration file XML DTD

The configuration file for the NDB parameters shall be formatted in XML format, complying to the following DTD.

```
<!-- DTD for Medspiration MDB Configuration file ================= -->

<!ELEMENT MDB_parameters (space_constraint, time_constraint)>


<!-- Entry space_constraint ============================================= -->
<!ELEMENT space_constraint (matchup_distance_km)>
<!ELEMENT matchup_distance_km (#PCDATA)>

<!-- Entry time_constraint ============================================= -->
<!ELEMENT time_constraint (matchup_time_hours)>

<!ELEMENT matchup_time_hours (#PCDATA)>
```

A9.7.2 MDB parameters configuration file XML sample file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE MDB_parameters SYSTEM "reference_parameters.dtd">
<MDB_parameters>
   <space_constraint>
      <matchup_distance_km>25</ matchup_distance_km >
   </space_constraint>
   < time_constraint >
      < matchup_time_hours >100</ matchup_time_hours >
   </ time_constraint >
</MDB_parameters>
```

Appendix A10 Matlab code for Stuart-Menteth Diurnal Variation Parameterisation

This Appendix provides information for groups implementing the diurnal variation parameterisation of Stuart Menteth.

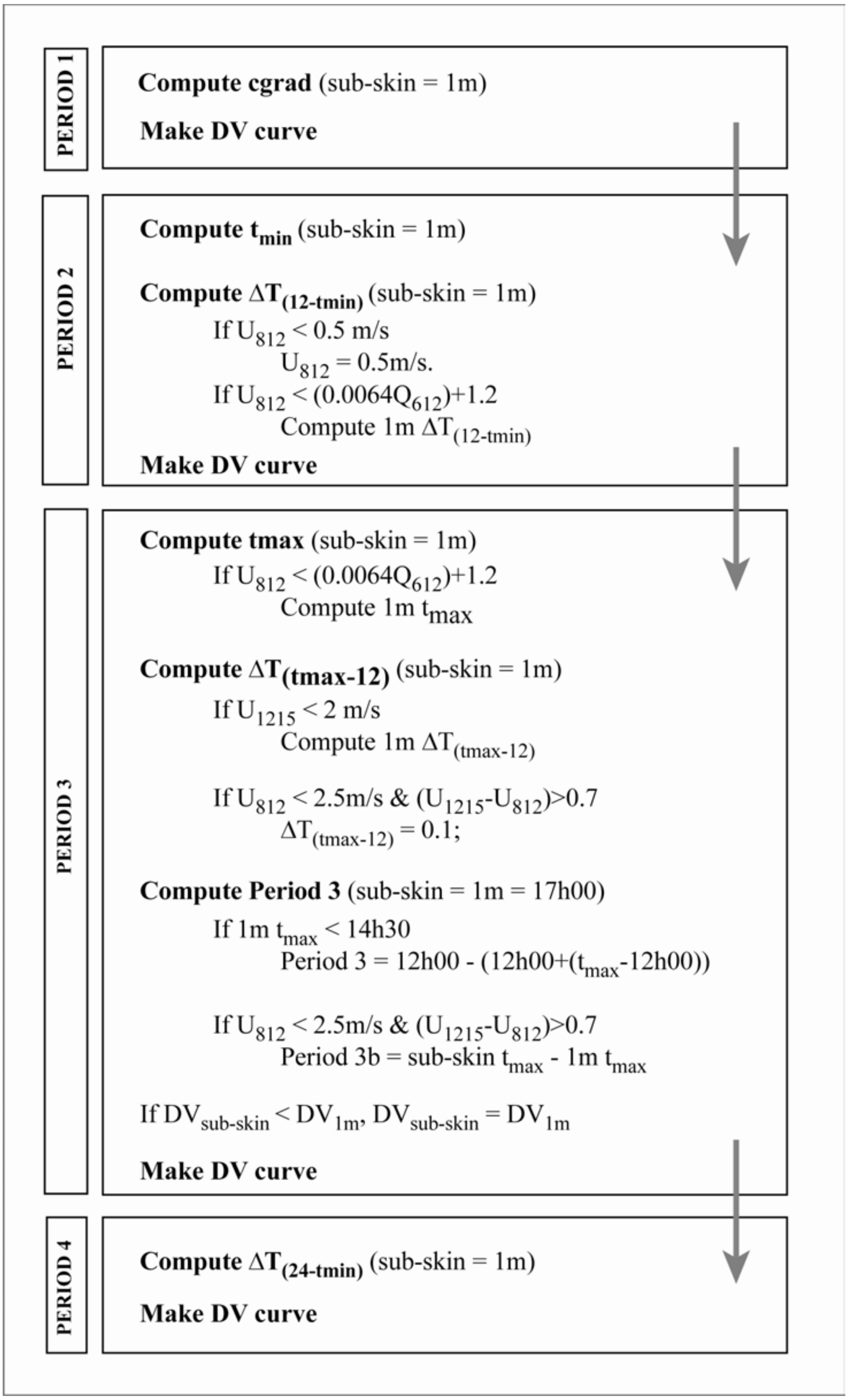


Figure A10.1 Flow diagram explaining how the Stuart-Menteth model in implemented (A. Stuart-Menteth)

The following MATLAB code implements the scheme shown in Figure A10.1.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Stuartmenteth_paramv2.m
% Version 2 (19/2/2004)
%
% code written by Alice Stuart-Menteth
% Sent to Pierre Le Borgne (Meteo France) 19/2/2004, S. Woolnough, Bruno
%****
% This routine estimates the shape (phase and magnitude) of the diurnal
% SST cycle for 1m and subskin using only wind and insolation measurements
% from different times of the day.
%
% Required INPUTS:
% dt = time step of parameterisation (hrs)
% wind and insolation measurements to compute
%  1. tu06: mean(wind 0h-6h) - or point measurement within that time period
%  2. tu812: mean(wind 8h-12h) - or point measurement within that time period
%  3. tu1215: mean(wind 12h-15h) - or point measurement within that time period
%  4. tu1624: mean(wind 16h-24h) - or point measurement within that time period
%  5. tq612: mean(qsw 6h-12h)
%  6. tq1218: mean(qsw 12h-18h)
%
% 6 models derived:
%  i)   cgrad (morning cooling gradient K/hr)
%  ii)  tmin (time of daily min SST DV=0)
%  iii) dt1s & dt11m (SST12h-SSTmin for sub-skin and 1m)
%  iv)  tmaxs & tmax1m (time of daily maximum SST for sub-skin and 1m)
%  v)   dt2s & dt21m (SSTmax-SST12 for sub-skin and 1m)
%  vi)  dt3 (SST24-SSTmin)
%
% OUTPUT:
% msub - sub-skin model of diurnal SST variation referenced from time of minimum SST where
msub=0;
% m1m  - 1m model of diurnal SST variation referenced from time of minimum SST where m1m=msub=0;
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%**INPUTS to be defined
% Time step of model (hrs) - can be altered to suit user requirements
dts=.25;

% wind and insolation (tu06, tu812, tu1216, tu1624, tq612, tq1218)
% insert files


%%%%%%%%%%%%%%%%% START OF MODEL %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Defining key constants
time=0.25:dts:24;
ii=length(time); %no. times steps over 24 hrs
jj=length(tu812);%no. days in data set

h24=ii; %Time step at 24hrs
h12=ii/2; %Time step at 12hrs
h17=17/24*ii; %Time step at 17hrs
m1m=NaN*ones(jj,ii);
msub=NaN*ones(jj,ii);


%*** TIME 1 - 00:00 to Tmin (morning cooling)  ****************************

%------- computing cgrad (morning cooling gradient)
c1=[0.0034 0.0486];
cgrad=c1(1)+c1(2).*exp(-0.3.*tu06);
neg=find(cgrad<0); cgrad(neg)=0; neg=[];
%---------------------------------------------------------------

%------- computing tmin (time of daily minimum SST)
c2=[6.4724 0.1766];
tmin=c2(1)+c2(2).*log(tu812);
```

```
%finding nearest time point to tmin
tt=floor(tmin*h24/24); %tmin=tt*dt;
%-----------------------------------------------------------------

%------- computing curve 1
for i=1:jj;
    t=tt(i);pmt1=time(1:t);
    m1m(i,1:t)=(cgrad(i).*tmin(i)).*((pmt1-tmin(i)).^2)./(((0-tmin(i)).^2));
    msub(i,1:t)=m1m(i,1:t);pmt1=[];
end
%-----------------------------------------------------------------
%*** End of Time 1 ********************************************************



%*** TIME 2 (morning heating) *********************************************

%------- computing dt1 (computing morning heating)
limit=find(tu812<0.5);tu812(limit)=0.5;
c3=[0.009 6.3071 0.0022 0.00018 0.0114];
dt1s=c3(1)+c3(2).*exp(-(600./tq612).*tu812)+c3(3).*tq612./tu812;
dt11m=dt1s;

%if top metre stratified in morning - 1m different from sub-skin, re-compute 1m
strat_am=find(tu812<((tq612.*0.0064)+1.2));

if(isempty(strat_am)==0)
    ustr=(tq612(strat_am).*0.0064)+1.2;
    dt11m(strat_am)=c3(1)+c3(2).*exp(-
ustr.*(600./tq612(strat_am)))+c3(3).*(tq612(strat_am)./ustr)+(((c3(4).*tq612(strat_am))+c3(5)).*
(ustr-tu812(strat_am)));

end
%--------------------------------------------------------------------

%------- computing curve 2
for i=1:jj; pmt2=time(tt(i):h12);
    msub(i,tt(i):h12)=(dt1s(i)./((12-tmin(i)).^2)).*((pmt2-tmin(i)).^2);
    m1m(i,tt(i):h12)=(dt11m(i)./((12-tmin(i)).^2)).*((pmt2-tmin(i)).^2);
end
%--------------------------------------------------------------------
%*** End of Time 2 **********************************************************



%*** TIME 3 - 12:00 to 17:00 (afternoon heating) ****************************

%------- computing tmax (time of daily maximum SST)
c4=[13.0157 -1.9736 2.2835 0.0042 -0.002];
tmaxs=c4(1)+c4(2).*log(tu1215)+c4(3).*log(tu812)+c4(4).*tq1218+c4(5).*tq612;
tmax1m=tmaxs;

%if top metre stratified in morning - 1m different from sub-skin, re-compute 1m
if(isempty(strat_am)==0)
 c4b=[15.75 -1.9109 .8754 0.0035 -0.0027];

tmax1m(strat_am)=c4b(1)+c4b(2).*log(tu1215(strat_am))+c4b(3).*log(tu812(strat_am))+c4b(4).*tq121
8(strat_am)+c4b(5).*tq612(strat_am);
end

mxt=floor(tmaxs*h24/24); % time step at sub-skin tmax
mxt1=floor(tmax1m*h24/24); % time step at 1m tmax
%----------------------------------------------------------------------


%------- computing dt2 & dt (afternoon heating and total diurnal SST amplitude)
limit=find(tu1215<0.5);tu1215(limit)=0.5;
c5=[-0.0541 2.057 0.002];

dt2s=c5(1)+c5(2).*exp(-(600./tq1218).*tu1215) +c5(3).*tq1218./tu1215;
dt21m=dt2s;

% if top metre stratified in afternoon - 1m different from sub-skin, re-compute 1m
strat_pm=find(tu1215<2);
if(isempty(strat_pm)==0)
    dt21m(strat_pm)=c5(1)+c5(2).*exp(-(600./tq1218(strat_pm)).*2) +c5(3).*tq1218(strat_pm)./2;
```

```
end

% if wind in afternoon breaks down strong morning stratification
break_strat=find(tu812<2.5 & (tu1215-tu812)>.7);dt2s(break_strat)=0.1;

%to keep realistic shape (rarely needed)
neg=find(dt2s<0.05);  dt2s(neg)=0.05; neg=[];
neg=find(dt21m<0.05);  dt21m(neg)=0.05; neg=[];

%computing daily maximum diurnal warming amplitude
dts=dt1s+dt2s;
dt1m=dt11m+dt21m;
%-------------------------------------------------------------------


%------- computing curve 3
 h17=h17*ones(1,jj);
    nn=find(tmax1m<14.5); h17(nn)=(round((tmax1m(nn)./24*96))-
(h12+1))+(round((tmax1m(nn)./24*96)))+1;

 for i=1:jj;
    pmt3=time(h12+1:h17(i));
    m1m(i,h12+1:h17(i))=dt1m(i)-(dt21m(i)./((12-tmax1m(i)).^2)).*((pmt3-tmax1m(i)).^2);
    msub(i,h12+1:h17(i))=dts(i)-(dt2s(i)./((12-tmaxs(i)).^2)).*((pmt3-tmaxs(i)).^2);

    %to keep realistic shape (rarely needed)
    negs=find(m1m(i,h12+1:h17(i))<0); m1m(i,negs+h12)=0;
    negs=find(msub(i,h12+1:h17(i))<0); msub(i,negs+h12)=0;
end


%considering days with slight increase in wind in pm: reduces sub-skin and increases 1m
if(isempty(break_strat)==0)
  for i=1:length(break_strat);

     if (tu812(break_strat(i))<2 & tq612(break_strat(i))>400)
      pmt3=time(mxt(break_strat(i))+1:mxt1(break_strat(i)));

msub(break_strat(i),mxt(break_strat(i))+1:mxt1(break_strat(i)))=dt1m(break_strat(i))+((dts(break
_strat(i))-dt1m(break_strat(i)))./(tmax1m(break_strat(i))-tmaxs(break_strat(i))).^2.*(pmt3-
tmax1m(break_strat(i))).^2);

msub(break_strat(i),mxt1(break_strat(i))+1:h17(break_strat(i)))=m1m(break_strat(i),mxt1(break_st
rat(i))+1:h17(break_strat(i)));
     end

  end
end
%-----------------------------------------------------------------

%-----------------------------------------------------------------
%** making sub-skin join 1m curve during afternoon cooling
%** if sub-skin DV < 1m DV, sub-skin = 1m
for j=1:jj; k=0;ty=mxt((j));join(j)=0;
    for i=ty:h17(j); g=msub((j),i)-m1m((j),i);
        if g<0; k=k+1; if(k==1); join(j)=i;end;end;
    end;

    if(join(j)>0 & join(j)<68)
      msub(j,join(j):h17(j))=m1m(j,join(j):h17(j));
      dg=(msub(j,join(j)-1)-msub(j,join(j)+1))./2; msub(j,join(j))=msub(j,join(j)-1)-dg;
    end
end
%---------------------------------------------------------------
%*** End of Time 3**********************************************************



%*** TIME 4 - 17:00 to 24:00 (night-time cooling) *************************

%------- computing dt3 (dt(24hr-tmin))
c6=[0.1365 0.209 -0.017];
dt31m=c6(1)+c6(2).*(dt1m)+c6(3).*tu1624;
neg=find(dt31m<0); dt31m(neg)=0;
dt3s=dt31m;
%-------------------------------------------------------------
```

```
%------- computing curve 4
for i=1:jj;
   pmt4=time(h17(i):h24);

   coamp(i)=m1m(i,h17(i))-dt31m(i)'; %1m
   % to keep realistic shape at high winds
   if(coamp(i)<0); coamp(i)=m1m(i,h17(i));dt31m(i)=0; dt3s(i)=0;end

   scoamp(i)=msub(i,h17(i))-dt3s(i)'; %sub-skin
   % to keep realistic shape at high winds
   if(scoamp<0); ;scoamp(i)=msub(i,h17(i));dt3s(i)=0; end

   m1m(i,h17(i):h24)=dt31m(i)+coamp(i)'./(((24-time(h17(i))).^2)).*((pmt4-24).^2);
   msub(i,h17(i):h24)=dt3s(i)+scoamp(i)'./(((24-time(h17(i))).^2)).*((pmt4-24).^2);

end
%-----------------------------------------------------------------
%****************************************************************************


%********** optional tidying up
% smoothing joins between curves 2 & 3 and curves 3 & 4
dg=(m1m(:,h12+1)-m1m(:,h12-1))./2; m1m(:,h12)=m1m(:,h12+1)-dg; m1m(:,h12-1)=m1m(:,h12+1)-dg.*2;
dg=[];dg=(msub(:,h12+1)-msub(:,h12-1))./2; msub(:,h12)=msub(:,h12+1)-dg; msub(:,h12-
1)=msub(:,h12+1)-dg.*2;

dg=[];dg=(m1m(:,h17-1)-m1m(:,h17+1))./2; m1m(:,h17)=m1m(:,h17-1)-dg;m1m(:,h17+1)=m1m(:,h17-1)-
dg*2;
dg=[];dg=(msub(:,h17-1)-msub(:,h17+1))./2; msub(:,h17)=msub(:,h17-1)-dg;
msub(:,h17+1)=msub(:,h17-1)-dg*2;
%******************************


%%%%%%%%%% END OF MODEL %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Published by the International GHRSST-PP Project Office
Met Office
Beagle B2-042
Fitzroy Road
Exeter, EX3 1PB
United Kingdom

Tel: +44 (0)1392 886622
Fax: +44 (0)1393 885681
E-mail: craig.donlon@metoffice.com

The GHRSST-PP International Project Office is sponsored by the European Space Agency and the MetOffice , United Kingdom.